

Fixed bandwidth inference for fractional cointegration*

Javier Hualde[†]

Universidad Pública de Navarra

Fabrizio Iacone

Università degli Studi di Milano and University of York

January 22, 2019

Abstract

In a fractional cointegration setting we derive the fixed bandwidth limiting theory of a class of estimators of the cointegrating parameter which are constructed as ratios of weighted periodogram averages. These estimators offer improved limiting properties over those of more standard approaches like OLS or NBLS estimation. These advantages have been justified by means of traditional asymptotic theory and here we explore whether these improvements still hold when considering the alternative fixed bandwidth theory and, more importantly, whether this latter approach provides a more accurate approximation to the sampling distribution of the corresponding test statistics. This appears to be relevant, especially in view of the typical oversizing displayed by Wald statistics when confronted to the standard limiting theory. A Monte Carlo study of finite-sample behaviour is included.

*We are grateful to Morten Ø. Nielsen and two anonymous referees for very useful comments and suggestions. We also thank participants at the VIII_t Workshop in Time Series Econometrics at the University of Zaragoza and at the Long Memory Conference at Aalborg University for helpful comments.

[†]Javier Hualde's research is supported by the Spanish Ministerio de Economía y Competitividad through project ECO2015-64330-P.

1 Introduction

Since the works of Kiefer and Vogelsang (2002, 2005), there has been a growing interest on an alternative approach to deriving asymptotic theory denoted as fixed- b asymptotics. The original idea was due to Neave (1970), who studied the limiting behaviour of standard nonparametric estimators of the spectral density. These estimators depend on several items, including the sample size n and a user-chosen number known as bandwidth (M), and they are consistent under standard conditions which include

$$\frac{1}{M} + \frac{M}{n} \rightarrow 0,$$

as $n \rightarrow \infty$. Alternatively, Neave (1970) studied the limiting properties of these estimators when $M/n \rightarrow b \in (0, 1]$ as $n \rightarrow \infty$. This has been termed fixed- b asymptotics (in contrast to letting $b = 0$, known as small- b asymptotics), which is motivated by the fact that in any practical situation a non-zero fraction M/n is used, so fixing the proportion M/n in the asymptotics could yield a better approximation to the sampling distribution of the estimator. These advantages have been illustrated by Kiefer and Vogelsang (2002, 2005), Bunzel and Vogelsang (2005) and Iacone, Leybourne and Taylor (2013). The theoretical reason behind this phenomenon has been studied by Jansson (2004) and Sun, Phillips and Jin (2008) who justified by means of higher order Edgeworth expansions that the fixed- b asymptotics lead to a more refined approximation than the traditional small- b approach for a Gaussian location model. Focusing also on this model, Sun (2014) provided a similar justification, which he also extended to a generalized method of moments setting.

In the somewhat different context of standard cointegration, where observables are unit roots and cointegrating errors are weakly dependent covariance stationary processes, Bunzel (2006) derived the fixed- b limit of a Wald test statistic based on the dynamic ordinary least squares (DOLS) estimator of the cointegration parameter and on a weighted covariance estimator of the long-run variance. Similarly, Jin, Phillips and Sun (2006) derived a fixed- b theory for tests based on the fully modified ordinary least squares (FM-OLS) estimator. Both results have been referred by Vogelsang and Wagner (2014) as “partial” fixed- b theory. The reason is that Bunzel’s (2006) analysis ignores the impact of lead and lag choices on the implementation of the DOLS (that is, the fixed- b theory is just considered for the nonparametric estimator of the long run variance), whereas Jin, Phillips and Sun’s (2006) approach requires standard consistency results, therefore ignoring the choice of tuning parameters inherent to the FM-OLS estimator. On the contrary, Vogelsang and Wagner (2014) derive the

“complete” fixed- b limit of the FM-OLS which, given its dependence on nuisance parameters, is not suitable for statistical inference. Thus, they propose an alternative estimator of the cointegrated parameter (integrated modified OLS), which does not depend on choices of tuning parameters and which is used to construct a Wald statistic whose fixed- b limiting distribution is pivotal.

In this paper we will explore a similar idea to fixed- b asymptotics in the more general fractional cointegration setting. In particular we will focus on the simple semiparametric model

$$y_t = \nu x_t + \Delta^{-\gamma} \{u_{1t}1(t > 0)\}, \quad (1)$$

$$x_t = \Delta^{-\delta} \{u_{2t}1(t > 0)\}, \quad (2)$$

for $t = 0, \pm 1, \dots$, where $u_t = (u_{1t}, u_{2t})'$ is a bivariate covariance stationary unobservable process with zero mean and nonparametric spectral density matrix $f(\lambda)$ that is at least nonsingular at frequency zero and continuous at all frequencies, $\Delta = 1 - L$, where L is the lag operator, the fractional difference operator is given formally, for any real α , $\alpha \neq -1, -2, \dots$, by

$$\Delta^{-\alpha} = \sum_{j=0}^{\infty} a_j(\alpha) L^j, \quad a_j(\alpha) = \frac{\Gamma(j + \alpha)}{\Gamma(\alpha)\Gamma(j + 1)},$$

with Γ denoting the gamma function, $1(\cdot)$ is the indicator function and $\delta > \gamma > -1/2$. When $\gamma = 0$, $\delta = 1$, this corresponds to the usual bivariate cointegrated $I(1)/I(0)$ system, but (1), (2) allows for many other situations covering stationary or nonstationary cointegration, where $\delta < 1/2$ or $\delta \geq 1/2$, respectively. Note that in this latter case the truncations in (1), (2) ensure that the processes y_t , x_t are well defined in mean square sense. Processes x_t and y_t are said to have integration order δ and are called $I(\delta)$, while the cointegrating error $\Delta^{-\gamma} \{u_{1t}1(t > 0)\}$ is $I(\gamma)$, so whenever $\gamma < \delta$, y_t and x_t are cointegrated. More detailed conditions will be imposed below.

For a slightly more general model than (1), (2), Hualde and Iacone (2015) examined the fixed- b approximation to the distribution of the weighted covariance (WC) estimator

$$\tilde{\nu} = \left(\sum_{l=-n+1}^{n-1} k(l/M) c_{xx}(l) \right)^{-1} \sum_{l=-n+1}^{n-1} k(l/M) c_{xy}(l), \quad (3)$$

where $k(x)$ is a kernel function and, for two generic sequences ξ_t , ζ_t , $c_{\xi\zeta}(l) = n^{-1} \sum_{t=1}^{n-l} \xi_t \zeta'_{t+l}$ for $l \geq 0$; $= n^{-1} \sum_{t=1-l}^n \xi_t \zeta'_{t+l}$ for $l < 0$, prime denoting transpo-

sition. In addition, they provided Monte Carlo evidence showing that the fixed- b approximation is more accurate than the traditional one which imposes (at least) $M = o(n)$, and which was derived by Phillips (1991a) for the standard $I(1)/I(0)$ case and by Marinucci (2000) for the fractional general setting.

Hualde and Iacone (2015) also pinpointed the connection between the fixed- b asymptotic theory for WC and the fixed bandwidth approach for narrow band least squares (NBLS) estimation. In particular, $\tilde{\nu}$ in (3) has a frequency domain representation given by

$$\tilde{\nu} = \left(\int_{-\pi}^{\pi} K_M(\lambda) I_{xx}(\lambda) d\lambda \right)^{-1} \int_{-\pi}^{\pi} K_M(\lambda) I_{xy}(\lambda) d\lambda,$$

where $K_M(\lambda) = (2\pi)^{-1} \sum_{|l| < n} k(l/M) e^{-il\lambda}$ is the spectral window associated to $k(\cdot)$ and for generic sequences ξ_t, ζ_t , define the discrete Fourier transform and (cross-) periodogram

$$w_{\xi}(\lambda) = \frac{1}{(2\pi n)^{\frac{1}{2}}} \sum_{t=1}^n \xi_t e^{it\lambda}, \quad I_{\xi\zeta}(\lambda) = w_{\xi}(\lambda) w_{\zeta}(-\lambda)', \quad I_{\xi}(\lambda) = I_{\xi\xi}(\lambda).$$

Then, approximating integrals by sums over the Fourier frequencies $\lambda_j = 2\pi j/n$ for $j = 0, \pm 1, \dots, \pm [n/2]$, where $[\cdot]$ means integer part, and using the Daniell kernel, we get an approximation to $\tilde{\nu}$ given by

$$\hat{\nu}_m = \left(\sum_{j=0}^m c_j I_{xx}(\lambda_j) \right)^{-1} \sum_{j=0}^m c_j \operatorname{Re} I_{xy}(\lambda_j), \quad (4)$$

where $c_j = 1, j = 0, n/2, c_j = 2$, otherwise and $m = [n/(2M)]$. In fact (4) is the NBLS estimator, whose limiting properties have been studied under the assumption

$$\frac{1}{m} + \frac{m}{n} \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (5)$$

Condition (5) implies that the estimator uses a degenerating band of frequencies around the origin, therefore focusing on the long run components of the observables, and, also, that it gathers increasing information by letting m grow with the sample size. Under (5) and additional regularity conditions, Robinson and Marinucci (2003) derived the rate of convergence of the NBLS for the (asymptotically) stationary case (with $\delta < 1/2$). For this stationary setting, imposing also $\gamma \geq 0, \gamma + \delta < 1/2$, Christensen and Nielsen (2006) derived a Gaussian limiting distribution for the NBLS

for the particular case of noncoherence between regressor and error at the zero frequency. Imposing also (5), the limiting properties of the NBLS for the nonstationary case (with $\delta \geq 1/2$) were derived by Robinson and Marinucci (2001).

Comparing (3) and (4), noting that $m = \lfloor n/(2M) \rfloor$, (5) implies that $M/n \rightarrow 0$. Then letting $M = bn$, $b > 0$ and using the Daniell kernel corresponds to discussing the properties of the NBLS when m is fixed ($m = \lfloor 1/(2b) \rfloor$). In fact, related to this possibility, Robinson and Marinucci (2001) conjectured that for the case $\gamma \geq 0$, $\gamma + \delta \leq 1$, $(\gamma, \delta) \neq (0, 1)$, the faster $n^{\delta-\gamma}$ convergence rate could be achievable by the NBLS with m fixed as $n \rightarrow \infty$. Chen and Hurvich (2003) verified this conjecture for a tapered NBLS applied to differenced data.

Hualde and Iacone (2017) exploited this idea, which they named as fixed bandwidth asymptotics, and applied it to the studentized mean of a covariance stationary fractionally integrated process where the long run variance was estimated by an averaged periodogram estimator (where the degree of averaging was kept fixed). This approach falls within the fixed-smoothing asymptotics setting, term coined by Zhang and Shao (2013), which characterizes situations where a smoothing parameter is held fixed in the asymptotics. This includes the number of basis functions in nonparametric estimation, see, e.g., Sun (2013), the number of non-overlapping blocks on which a subsampling based t -statistic is constructed, see, e.g., Ibragimov and Müller (2010) and Zhang and Shao (2013), or the number of cluster groups, see, e.g., Bester, Conley and Hansen (2011).

In this vein, we derive in the present paper the fixed bandwidth (hereinafter denoted as fixed- m) limiting theory of various estimators of ν . First, we will obtain the fixed- m limit of the NBLS estimator (4). As with the large- m limiting theory (with $m \rightarrow \infty$ as $n \rightarrow \infty$), the fixed- m limit of $\hat{\nu}_m$ depends on nuisance parameters, so, except in some specific cases (see, e.g., Christensen and Nielsen, 2006), it is unsuitable for statistical inference.

Some improvements over NBLS estimation have been proposed. First, for the stationary case, Nielsen (2005) considered a general class of estimators which include the NBLS and a narrow band generalized least squares (NBGLS) as special cases. Under the condition of noncoherence at frequency zero between regressors and errors, Nielsen (2005) showed that, as expected, the NBGLS is more efficient than NBLS. Nielsen's (2005) approach could be denoted more precisely as subsystem NBGLS: for the particular setting in (1), (2), this estimator is inspired by the GLS applied to the first equation (1). There is, therefore, room for improvement by considering a full system NBGLS taking also into account the structure of (2). Actually, this

discussion is the semiparametric equivalent to the distinction between full system maximum likelihood (ML) and subsystem limited information ML estimation (see, Phillips, 1991b). As will be seen below, our proposal can be viewed as a full system NBGLS.

As an attempt to relax the strong noncoherence condition, Nielsen and Frederiksen (2011) proposed for the weak cointegration case (with $\delta - \gamma < 1/2$) a modification of the NBLs, termed Fully-modified NBLs, which eliminates the bias due to nonzero coherence at the origin and which retains, as in the case of zero coherence, the same convergence rate and also the asymptotic normality.

Alternatively, we will analyze a class of estimators of the cointegrating parameter in model (1), (2) which are constructed as ratios of periodogram averages weighted by multiplicative factors based on the inverse of smoothed estimators of $f(0)$. Through this weighting, these estimators also eliminate the endogeneity bias, therefore offering important advantages over the alternative NBLs estimation: their convergence rate is never slower than that of NBLs and, under standard conditions (which traditionally have included a bandwidth growing with the sample size), lead to Wald test statistics with the same standard null $N(0, 1)$ limit distribution irrespective of whether there is “strong” or “weak” cointegration, where $\delta - \gamma > 1/2$ or $\delta - \gamma < 1/2$, respectively. These results were formally justified in the large- m framework by Hualde and Robinson (2010). Related results for different contexts include Robinson and Hualde (2003), Johansen and Nielsen (2012) and Andersen and Varneskov (2018). Within this setting, our main aim in the present paper is to explore whether these nice limiting properties still hold when considering the alternative fixed- m theory and, more importantly, whether this approach provides a more accurate approximation to the sampling distribution of the corresponding test statistics. This appears to be relevant, especially in view of the typical oversizing displayed by Wald statistics when confronted to the standard limiting theory.

The following section presents the fixed- m limit of the NBLs estimator. Next, in Sections 3 and 4, we will introduce two levels of analysis. First, we discuss a “partial” fixed- m theory, where the bandwidth characterizing the estimator of the cointegrating parameters is kept fixed, but we employ consistency arguments for the estimators of the nuisance parameters (and therefore assume here increasing bandwidth). Next, in Section 4 we will present a more “complete” fixed- m theory, which captures the effect of the bandwidth choice in the estimators of some nuisance parameters. Section 5 contains a Monte Carlo study of finite-sample behaviour. Unless otherwise stated, proofs are relegated to the Appendix.

2 Fixed- m asymptotic theory for NBLS estimation

We initially derive the fixed- m limit of the NBLS estimator (4). As anticipated, alternative estimators of ν enjoy improved properties over NBLS, but, nevertheless, we include $\widehat{\nu}_m$ in our fixed- m analysis mainly for completeness. Incidentally, our results verify Robinson and Marinucci's (2001) conjecture, whose large- m limiting theory can be found in their Propositions 6.1-6.3 and 6.5. Note, however, that these results were just given for the case of nonstationary x_t , i.e., $\delta \geq 1/2$, and $\gamma \geq 0$. On the contrary, our fixed- m results will cover all situations as long as $\delta > \gamma > -1/2$.

We impose the following regularity condition. Let I_a be the a -rowed identity matrix and $\beta = \delta - \gamma$, so β denotes the strength of the cointegration relation.

Assumption 1. *The process u_t , $t = 0, \pm 1, \dots$, has representation*

$$u_t = A(L) \varepsilon_t, \quad A(z) = I_2 + \sum_{j=1}^{\infty} A_j z^j,$$

where $\det \{A(1)\} \neq 0$, $A(e^{i\lambda})$ is differentiable in $\lambda \in [-\pi, \pi]$ with derivative in $Lip(\eta)$, $\eta > 1/2$; with $\|\cdot\|$ denoting the Euclidean norm, ε_t is a $(2p)$ th-order stationary process with bounded $(2p)$ th-order spectral density such that $p \geq 2$ and if $\beta > 1/2$, $p > 1/(2\beta - 1)$; defining $E(\varepsilon_t \varepsilon_t') = \Sigma$, where Σ is positive definite,

$$E(\varepsilon_t | \mathcal{F}_{t-1}) = 0, \quad E(\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}) = \Sigma$$

almost surely, where F_t is the σ -field of events generated by ε_s , $s \leq t$. Also, conditional (on F_{t-1}) third and fourth moments of ε_t equal the corresponding unconditional moments (almost surely).

Assumption 1 is satisfied if u_t is a stationary autoregressive-moving average (ARMA) process. It avoids requiring independence or identity of distribution of ε_t (like it is imposed in Hualde and Robinson, 2010), but rules out conditional heteroskedasticity. With the exception of the strong moment condition, which is required to justify the use of the fractional invariance principle of Hosoya (2005), this condition is standard in the time series literature since Hannan (1973) and it is very similar to that in Nielsen and Frederiksen (2011). Note that the strong moment condition (when β is above but close to $1/2$) is satisfied for Gaussian ε_t . This assumption also imposes a global smoothness condition on $f(\lambda)$ which implies that

$$\sum_{j=1}^{\infty} j \|A_j\| < \infty, \quad \sum_{r=-\infty}^{\infty} |r| \|\Gamma(r)\| < \infty,$$

where $\Gamma(r) = E(u_t u'_{t-r})$.

We introduce further notation. Let $\zeta = (1, 0)'$, $\xi = (0, 1)'$, $B(r)$ be a Brownian motion with covariance matrix Σ , and for $d > 1/2$,

$$B(r; d) = \int_0^r \frac{(r-s)^{d-1}}{\Gamma(\beta)} dB(s), \quad S_\alpha(r; d) = \alpha' A(1) B(r; d), \quad \alpha = \zeta, \xi,$$

so $B(r; d)$ is a Type II fractional Brownian motion. Finally, denote by \rightarrow_d convergence in distribution, and, for any functional $G(r)$, let $\widehat{G}(r) = G(r) - rG(1)$ and let the long run variance of u_t be $\Omega = 2\pi f(0)$.

Theorem 1. *Let Assumption 1 hold and m be any nonnegative fixed number. Then, as $n \rightarrow \infty$,*

$$n^\beta(\widehat{\nu}_m - \nu) \rightarrow_d \left(A_m^{\xi, \xi}(\Omega, \delta, \delta) \right)^{-1} A_m^{\zeta, \xi}(\Omega, \gamma, \delta), \quad (6)$$

where

$$\begin{aligned} A_m^{\alpha_1, \alpha_2}(\Omega, d_1, d_2) = & \left\{ S_{\alpha_1}(1; d_1 + 1) S_{\alpha_2}(1; d_2 + 1) + 2 \sum_{j=1}^m (2\pi j)^2 \right. \\ & \times \left[\int_0^1 \sin(2\pi j r) \widehat{S}_{\alpha_1}(r; d_1 + 1) dr \int_0^1 \sin(2\pi j s) \widehat{S}_{\alpha_2}(s; d_2 + 1) ds \right. \\ & \left. \left. + \int_0^1 \cos(2\pi j r) \widehat{S}_{\alpha_1}(r; d_1 + 1) dr \int_0^1 \cos(2\pi j s) \widehat{S}_{\alpha_2}(s; d_2 + 1) ds \right] \right\}, \end{aligned}$$

taking the summation in $A_m^{\alpha_1, \alpha_2}(\Omega, d_1, d_2)$ as 0 whenever $m = 0$.

As mentioned before, Theorem 1 verifies Robinson and Marinucci's (2001) conjecture that the rate n^β is achievable with m fixed in the nonstationary case. For the stationary case, keeping m fixed implies in general a gain in convergence rate (from $(n/m)^\beta$ to n^β), although in the particular case studied by Christensen and Nielsen (2006) the faster rate $n^\beta m^{1/2-\beta}$ is achievable. Note that in (6), unlike the analysis of Christensen and Nielsen (2006), both numerator and denominator of the normalized and centered NBLS converge to random limits: these are the sum of crossproducts of Gaussian random variables corresponding to each of the Fourier transforms which conform the estimator (evaluated at the different Fourier frequencies). In general, the limiting distribution in (6) is nonstandard, but under the noncoherence condition of Christensen and Nielsen (2006) (which in our setting corresponds to a diagonal Ω) it is a mixture of Gaussians. In any case, comparing (6) with the asymptotic distribution derived under the large- m condition is not a very productive exercise: for a given m the NBLS is unique, but two different limiting approximations to its

sampling distribution can be obtained (large- m and fixed- m). Our claim is that the fixed- m limit is closer to the true sampling distribution and, as will be seen below, this is the main motivation of our analysis in the following sections.

3 “Partial” fixed- m asymptotic theory

In this section, we focus on alternative generalized least squares-type estimators which offer important advantages over NBLS estimation. Define

$$z_t(\gamma, \delta) = (y_t(\gamma), x_t(\delta))',$$

where for any sequence $\{\xi_t\}$, and real c , $\xi_t(c) = \Delta^c \{\xi_t 1(t > 0)\}$. Thus (1), (2) can be written as

$$z_t(\gamma, \delta) = \zeta x_t(\gamma) \nu + u_t 1(t > 0).$$

Define also

$$p(\lambda) = \zeta' f(\lambda)^{-1}, \quad q(\lambda) = \zeta' f(\lambda)^{-1} \zeta,$$

and letting Ω have ij th element ω_{ij} , set $\rho = \omega_{12}/\omega_{22}$ and $\omega_{1.2} = \omega_{11} - (\omega_{12}^2/\omega_{22})$. Then the infeasible bivariate version of Hualde and Robinson’s (2010) “zero-frequency” estimator is given by

$$\bar{\nu}_m(\gamma, \delta, \rho) = \frac{a_m(\gamma, \delta, \rho, \omega_{1.2})}{b_m(\gamma, \omega_{1.2})}, \quad (7)$$

where

$$\begin{aligned} a_m(\gamma, \delta, \rho, \omega_{1.2}) &= p(0) \operatorname{Re} \left\{ \sum_{j=0}^m c_j I_{z(\gamma, \delta)x(\gamma)}(\lambda_j) \right\}, \\ b_m(\gamma, \omega_{1.2}) &= q(0) \sum_{j=0}^m c_j I_{x(\gamma)}(\lambda_j), \end{aligned}$$

and $m \leq n/2$. Note that

$$q(0) = \frac{2\pi}{\omega_{1.2}}, \quad p(0) = \frac{2\pi}{\omega_{1.2}} (1, -\rho),$$

so the dependence of the estimator of ν on Ω only occurs through ρ . The infeasibility of $\bar{\nu}_m(\gamma, \delta, \rho)$ is due to its dependence on ρ , but also on the integration orders γ , δ . These are considered as known in most standard cointegration analyses (where $\gamma = 0$, $\delta = 1$), but in the more general fractional setting they are typically treated as unknown nuisance parameters.

With the aim of focusing solely on the impact of the assumptions over m on the limiting behaviour of $\bar{\nu}_m(\gamma, \delta, \rho)$ we will initially act as if γ, δ, ρ , were known.

Remark 1. For simplicity we have omitted any deterministic component in our model (1), (2). However, the possibility of nonzero deterministic components can be accommodated as in Robinson and Iacone's (2005) Model III. Allowing for deterministic terms would lead to a different estimator: this would be the semiparametric version of the parametric estimator given in (94), (95) of Robinson and Iacone (2005). The limiting distribution of this estimator can be easily analyzed by their arguments and the results below.

Remark 2. For a general multivariate case, the limiting properties of (7) under the traditional large- m condition have been studied by Hualde and Robinson (2010). Letting $f_{ij}(\lambda), f^{ij}(\lambda)$ be the (i, j) th components of $f(\lambda), f^{-1}(\lambda)$ respectively, the result is the following: under Assumption 1

(i) When $\beta > 1/2$, if

$$m^{-1} + m/n^\beta \rightarrow 0, \quad (8)$$

as $n \rightarrow \infty$, then

$$n^\beta(\bar{\nu}_m(\gamma, \delta, \rho) - \nu) \rightarrow_d \left\{ q(0) \int_0^1 S_\xi(r; \beta)^2 dr \right\}^{-1} 2\pi \zeta' A(1)^{-1'} \Sigma^{-1} \int_0^1 S_\xi(r; \beta) dB(r); \quad (9)$$

(ii) When $\beta < 1/2$, if

$$m^{\beta-1/2} \log^{1/2} n + m^{3+2\eta}/n^{2+2\eta} \rightarrow 0, \quad (10)$$

as $n \rightarrow \infty$, then

$$m^{\frac{1}{2}} \lambda_m^{-\beta}(\bar{\nu}_m(\gamma, \delta, \rho) - \nu) \rightarrow_d N\left(0, \frac{1 - 2\beta}{2f^{11}(0)f^{22}(0)}\right). \quad (11)$$

To be precise, Hualde and Robinson's (2010) Theorem 1 covers (9) and (11) for the case where ε_t is an independent and identically distributed sequence. Justifying these results under our less restrictive Assumption 1 is simple and requires the use of Hosoya's (2005) fractional invariance principle, noting also that the convergence to the stochastic integral in (9) follows by Theorem 2.2 of Kurtz and Protter (1991) (which is essentially the same result as in Jakubowski, Mémén and Pages, 1989).

Remark 3. Interestingly, the limit in (9) is a mixed normal random variable which, along with (11), can be exploited to construct the Wald test statistic

$$\bar{t}_m(\gamma, \delta, \rho, \omega_{1.2}) = b_m^{1/2}(\gamma, \omega_{1.2}) (\bar{\nu}_m(\gamma, \delta, \rho) - \nu)$$

with pivotal null asymptotic distribution. Hence, we also have that under the conditions of Theorem 1, for both $\beta > 1/2$ and $\beta < 1/2$, as $n \rightarrow \infty$,

$$\bar{t}_m(\gamma, \delta, \rho, \omega_{1.2}) \rightarrow_d N(0, 1),$$

which is a bivariate particularization of Hualde and Robinson's (2010) Corollary 1. Like the optimal procedures proposed for the standard $I(1)/I(0)$ cointegration setting (see, e.g., Johansen, 1988, 1991, Phillips, 1991a,b) or in the more general fractional framework (see Robinson and Hualde, 2003, Johansen and Nielsen, 2012), $\bar{t}_m(\gamma, \delta, \rho, \omega_{1.2})$ enjoys standard limit theory, and this is irrespective of whether there is weak or strong cointegration.

Given these large- m results, the question of interest is whether these results also apply whenever m is kept fixed. These are given in Theorem 2 and Corollary 1 below.

Theorem 2. *Let Assumption 1 hold and m be any nonnegative fixed number. Then, as $n \rightarrow \infty$,*

$$n^\beta (\bar{\nu}_m(\gamma, \delta, \rho) - \nu) \rightarrow_d D_m^{-1}(\Omega, \beta) N_m(\Omega, \beta), \quad (12)$$

where $D_m(\Omega, \beta) = \zeta' \Omega^{-1} \zeta A_m^{\xi, \xi}(\Omega, \beta, \beta)$ and

$$\begin{aligned} N_m(\Omega, \beta) = & \zeta' \Omega^{-1} \left\{ S_\xi(1; \beta + 1) A(1) B(1) + 2 \sum_{j=1}^m (2\pi j)^2 \right. \\ & \times \left[\int_0^1 \sin(2\pi j r) \widehat{S}_\xi(r; \beta + 1) dr \int_0^1 \sin(2\pi j s) A(1) \widehat{B}(s) ds \right. \\ & \left. \left. + \int_0^1 \cos(2\pi j r) \widehat{S}_\xi(r; \beta + 1) dr \int_0^1 \cos(2\pi j s) A(1) \widehat{B}(s) ds \right] \right\}, \end{aligned}$$

taking the summations in $D_m(\Omega, \beta)$, $N_m(\Omega, \beta)$ as 0 whenever $m = 0$.

The proof is almost identical to that of Theorem 1, so it is omitted. The limiting distribution in (12) has a similar structure to that in (6), incorporating contributions from Fourier transforms at different Fourier frequencies, but with a key difference: for any $r, s \in [0, 1]$, $Cov(S_\xi(r; \beta + 1), \zeta' \Omega^{-1} A(1) B(s)) = 0$, so the limit in (12) is a mixed normal random variable.

The result in (12) leads to the following result.

Corollary 1. *Under the conditions of Theorem 2, as $n \rightarrow \infty$,*

$$\bar{t}_m(\gamma, \delta, \rho, \omega_{1.2}) \rightarrow_d N(0, 1).$$

Remark 4. The main implication of Corollary 1 is that, along with the result in Remark 3, the first order limiting distribution of $\bar{t}_m(\gamma, \delta, \rho, \omega_{1.2})$ is $N(0, 1)$ irrespective of the value of m . Thus, the (likely) dependence of the sampling distribution of $\bar{t}_m(\gamma, \delta, \rho, \omega_{1.2})$ on m vanishes when considering its first order approximation. This opens the door to considering a more refined approximation to clarify the way in which the dependence on m arises in subsequent smaller order terms of the asymptotic expansion. Undoubtedly, this would lead to a better understanding of the accuracy of the $N(0, 1)$ limit approximation for different values of m . In the simple setting of a Gaussian location model, Zhang and Shao (2013) derived a higher order Edgeworth expansion for the finite sample distribution of a subsampling-based t -statistic under a fixed-smoothing paradigm. In particular, the sample size was divided into K equal sized groups and the smoothing parameter K is held fixed, unlike with the standard procedure of considering a smoothing parameter which grows with respect to the sample size. Interestingly, Zhang and Shao (2013) showed that as K increases the first-order asymptotic approximation deteriorates. A formal analysis along these lines is beyond the scope of the present paper, but this issue will be explored by means of a Monte Carlo experiment in Section 5.

Remark 5. Incidentally, unlike the results presented in Remarks 2 and 3, Theorem 2 and Corollary 1 cover the somewhat elusive case $\beta = 1/2$, which is at the boundary between strong and weak cointegration (see, e.g., Robinson and Hualde, 2003 and Hualde and Robinson, 2007).

Next we discuss the behaviour of the feasible $\bar{\nu}_m(\hat{\gamma}, \hat{\delta}, \hat{\rho})$ for suitable estimators $\hat{\gamma}$, $\hat{\delta}$, $\hat{\rho}$ of γ , δ , ρ , respectively, for which the following conditions hold.

Assumption 2. *There exist $K < \infty$ and $\kappa > 0$ such that*

$$|\hat{\gamma}| + |\hat{\delta}| \leq K, \tag{13}$$

and

$$\hat{\gamma} = \gamma + O_p(n^{-\kappa}), \quad \hat{\delta} = \delta + O_p(n^{-\kappa}). \tag{14}$$

Condition (13) is innocuous if our estimators of the memory parameters optimize

over compact sets, as it occurs when using implicitly defined estimates like the local Whittle. This condition has been routinely applied in related settings, see, e.g., Robinson and Hualde (2003), Hualde and Robinson (2010).

Assumption 3. *There exists $\varkappa > 0$ such that*

$$\widehat{\rho} = \rho + O_p(n^{-\varkappa}). \quad (15)$$

Again, Assumption 3 is unprimitive, but it is very mild, even though the typical semiparametric estimators of the long run variance (and also of the integration orders) have a convergence rate no better than $n^{2/5}$. A similar condition has been employed by, e.g., Hualde and Robinson (2010) where it is discussed in detail.

Theorem 3 collects the results for $\bar{\nu}_m(\widehat{\gamma}, \widehat{\delta}, \widehat{\rho})$, noting that part (i) can be derived straightforwardly from the results in Hualde and Robinson (2010).

Theorem 3. *Let Assumptions 1-3 hold.*

(i) *Under the (8), (10) and*

$$(n^{-\kappa} + n^{-\varkappa}) m^{1-\max\{\min\{\beta, 1\}, 1/2\}} \log m \rightarrow 0, \quad (16)$$

as $n \rightarrow \infty$, results (9), (11) apply for $\bar{\nu}_m(\widehat{\gamma}, \widehat{\delta}, \widehat{\rho})$;

(ii) *If m is a nonnegative fixed number, as $n \rightarrow \infty$, result (12) applies for $\bar{\nu}_m(\widehat{\gamma}, \widehat{\delta}, \widehat{\rho})$.*

Remark 6. Condition (16) involves the convergence rates given in (14), (15), and letting $m \rightarrow \infty$ (as implied by (8), (10)), might not be straightforwardly satisfied given the previously mentioned bound for the rate of convergence of standard semiparametric estimators. For example, if m grows at the same rate as n , (16) does not hold when $\beta \leq 3/5$. Thus for moderate values of β either the rate of growth of m has to be restricted or the use of “improved” estimators of the nuisance parameters based on bias-reducing techniques might be required. For a fully detailed explanation about the complex interaction between the rates in (14), (15) and condition (16) see Hualde and Robinson (2010), pp. 496, 497. Alternatively, when keeping m fixed, just the mild conditions (14), (15) with $\kappa > 0$, $\varkappa > 0$ are needed.

Note that, even in the fixed- m case, Theorem 3 relies on consistency arguments for all estimators of the nuisance parameters. Imposing consistency of $\widehat{\gamma}$, $\widehat{\delta}$, seems unavoidable, but considering a fixed- m analysis for the estimators of functionals of Ω seems possible. In this sense the results in Theorem 3 (ii) are denoted as “partial”

fixed- m theory, in contrast with the more “complete” theory which will be addressed in Section 4.

Finally, we present the corresponding results to the feasible Wald statistic $\bar{t}_m(\hat{\gamma}, \hat{\delta}, \hat{\rho}, \hat{\omega}_{1.2})$, which requires an extra condition related to an estimator of $\omega_{1.2}$, say $\hat{\omega}_{1.2}$ (the same comment as that below Assumption 3 applies).

Assumption 4. *There exists $\varkappa > 0$ such that*

$$\hat{\omega}_{1.2} = \omega_{1.2} + O_p(n^{-\varkappa}).$$

Corollary 2. *Under the conditions of Theorem 3 and Assumption 4, as $n \rightarrow \infty$,*

$$\bar{t}_m(\hat{\gamma}, \hat{\delta}, \hat{\rho}, \hat{\omega}_{1.2}) \rightarrow_d N(0, 1).$$

This result can be justified as part (ii) of Theorem 3, so its proof is omitted.

4 “Complete” fixed- m asymptotic theory

As anticipated, in this section we derive the fixed- m limit of a feasible estimator of ν (and corresponding test statistic) without relying on consistency arguments for estimators of functionals of Ω (although we will still require consistency of $\hat{\gamma}, \hat{\delta}$). First, we will analyze the behaviour of $\bar{\nu}_m(\gamma, \delta, \tilde{\rho})$, for a particular estimator $\tilde{\rho}$ which will be specified below. The estimator $\bar{\nu}_m(\gamma, \delta, \tilde{\rho})$ is motivated as follows. Rewrite (1) as

$$y_t(\gamma) = \nu x_t(\gamma) + \rho x_t(\delta) + u_{1.2t}, \quad (17)$$

where $u_{1.2t} = u_{1t} - \rho u_{2t}$. Enlarging a regression model by introducing an additional regressor ($x_t(\delta)$) corresponds to a type of orthogonalization which combined with OLS estimation leads to an estimator of the cointegrating parameter with a standard limiting distribution in the white noise case (see, e.g., Phillips 1991b, Robinson and Hualde, 2003, Hualde and Robinson, 2007). In more general cases, a more sophisticated orthogonalization is required, like that implied by, e.g., the DOLS method. In contrast, as shown below, within the context of fixed- m asymptotics, the simple orthogonalization given by (17) leads to an estimator from which an asymptotically pivotal Wald statistic can be constructed even for the general u_t permitted by Assumption 1.

Setting

$$r_t(\gamma, \delta) = (x_t(\gamma), x_t(\delta))',$$

we define estimators

$$\begin{pmatrix} \hat{\nu}_m(\gamma, \delta) \\ \hat{\rho}_m(\gamma, \delta) \end{pmatrix} = \left(\operatorname{Re} \left\{ \sum_{j=0}^m c_j I_{r(\gamma, \delta)}(\lambda_j) \right\} \right)^{-1} \operatorname{Re} \left\{ \sum_{j=0}^m c_j I_{r(\gamma, \delta) y(\gamma)}(\lambda_j) \right\}, \quad (18)$$

which, strictly speaking, are infeasible due to their dependence on γ, δ . Note that $\hat{\nu}_m(\gamma, \delta)$ is not only the NBLS in the extended regression model (17), but it is also equal to $\bar{\nu}_m(\gamma, \delta, \tilde{\rho})$, where $\tilde{\rho} = \tilde{\omega}_{12}/\tilde{\omega}_{22}$, with

$$\tilde{\omega}_{12} = \frac{2\pi}{2m+1} \operatorname{Re} \left\{ \sum_{j=0}^m c_j I_{\hat{u}_1(\gamma, \delta) x(\delta)}(\lambda_j) \right\}, \quad \tilde{\omega}_{22} = \frac{2\pi}{2m+1} \sum_{j=0}^m c_j I_{x(\delta)}(\lambda_j), \quad (19)$$

and $\tilde{\omega}_{12}$ is constructed using residuals

$$\hat{u}_{1t}(\gamma, \delta) = y_t(\gamma) - \hat{\nu}_m(\gamma, \delta) x_t(\gamma).$$

The estimator $\hat{\nu}_m(\gamma, \delta)$ is particularly adequate to develop the “complete” fixed- m asymptotic theory which does not rely on assuming consistency for the estimator of ρ . Before presenting our main results we introduce further notation. Define $F(r; \beta) = (S_\xi(r; \beta), \xi' A(1) B(r))'$, $E(r) = (1, -\rho) A(1) B(r)$,

$$\hat{\Phi}_{\sin}(j) = \int_0^1 \sin(2\pi jr) \hat{F}(r; \beta + 1) dr, \quad \hat{\Phi}_{\cos}(j) = \int_0^1 \cos(2\pi jr) \hat{F}(r; \beta + 1) dr,$$

and for $m \geq 0$,

$$\begin{aligned} \Phi_m(\Omega, \beta) &= F(1; \beta + 1) F'(1; \beta + 1) + 2 \sum_{j=1}^m (2\pi j)^2 \left(\hat{\Phi}_{\sin}(j) \hat{\Phi}'_{\sin}(j) + \hat{\Phi}_{\cos}(j) \hat{\Phi}'_{\cos}(j) \right), \\ \phi_m(r; \Omega, \beta) &= F(1; \beta + 1) + 2 \sum_{j=1}^m 2\pi j \left(\hat{\Phi}_{\sin}(j) \cos(2\pi jr) - \hat{\Phi}_{\cos}(j) \sin(2\pi jr) \right), \end{aligned}$$

where the summations in $\Phi_m(\Omega, \beta)$, $\phi_m(r; \Omega, \beta)$ are 0 whenever $m = 0$. The fixed- m limit of the estimator in (18) is given in the next theorem.

Theorem 4. *Let Assumption 1 hold and m be any nonnegative fixed number. Then,*

as $n \rightarrow \infty$,

$$\begin{pmatrix} n^\beta (\hat{\nu}_m(\gamma, \delta) - \nu) \\ \hat{\rho}_m(\gamma, \delta) - \rho \end{pmatrix} \rightarrow_d \Phi_m^{-1}(\Omega, \beta) \int_0^1 \phi_m(r; \Omega, \beta) dE(r). \quad (20)$$

Remark 7. The limiting distribution of $\hat{\nu}_m(\gamma, \delta)$ implied by (20) has a similar structure to that in (12), but it incorporates the effect of estimating (inconsistently) the unknown parameter ρ . Nicely, this does not lead to a convergence rate loss for $\hat{\nu}_m(\gamma, \delta)$. Also, the limit in (20) is a mixed normal random variable because for any $r, s \in [0, 1]$, $Cov(F(r; \beta + 1), E(s)) = 0$, so it is expected that a Wald test statistic based on $\hat{\nu}_m(\gamma, \delta)$ enjoys standard asymptotics. Thus we define

$$\hat{t}_m(\gamma, \delta, \omega_{1.2}) = \frac{\hat{\nu}_m(\gamma, \delta) - \nu}{\left[\frac{\omega_{1.2}}{2\pi} \zeta' \left(\operatorname{Re} \left\{ \sum_{j=0}^m c_j I_{r(\gamma, \delta)}(\lambda_j) \right\} \right)^{-1} \zeta \right]^{\frac{1}{2}}},$$

noting that $\omega_{1.2}$ is also the variance of the Brownian motion $E(r)$ and that, apart from its dependence on γ, δ , $\hat{t}_m(\gamma, \delta, \omega_{1.2})$ is infeasible because $\omega_{1.2}$ is unknown. However, it is “more feasible” than $\bar{t}_m(\gamma, \delta, \rho, \omega_{1.2})$ in the sense that it incorporates an estimator of ρ . The next corollary presents the fixed- m limit of $\hat{t}_m(\gamma, \delta, \omega_{1.2})$.

Corollary 3. *Under the conditions of Theorem 4, as $n \rightarrow \infty$,*

$$\hat{t}_m(\gamma, \delta, \omega_{1.2}) \rightarrow_d N(0, 1).$$

Corollary 3 is similar to Corollary 1 although, interestingly, the $N(0, 1)$ limit is preserved even if ρ is not consistently estimated (see (20)). This phenomenon does not occur with respect to $\omega_{1.2}$: as will seen below, the lack of consistency of the estimator of $\omega_{1.2}$ plays a very relevant role in the derivation of the asymptotic behaviour of the corresponding feasible statistic.

The key ingredient of the feasible version of $\hat{t}_m(\gamma, \delta, \omega_{1.2})$ is the estimator of $\omega_{1.2}$. A very natural approach to estimating $\omega_{1.2}$ is to define residuals

$$\tilde{u}_{1.2t} = y_t(\gamma) - \hat{\nu}_m(\gamma, \delta) x_t(\gamma) - \hat{\rho}_m(\gamma, \delta) x_t(\delta),$$

and, mimicking (19), use the simple estimator

$$\widehat{\omega}_{1.2} = \frac{2\pi}{2m^* + 1} \sum_{j=0}^{m^*} c_j I_{\widehat{\omega}_{1.2}}(\lambda_j), \quad (21)$$

where the bandwidth m^* could be different from that employed in the estimation of ν and ρ . For reasons which will be apparent below the only requirement that this new bandwidth needs to fulfill is that $m^* \leq m$. Then, we define

$$\widehat{t}_m(\gamma, \delta, \widehat{\omega}_{1.2}) = \frac{\widehat{\nu}_m(\gamma, \delta) - \nu}{\left[\frac{\widehat{\omega}_{1.2}}{2\pi} \zeta' \left(\operatorname{Re} \left\{ \sum_{j=0}^m c_j I_{r(\gamma, \delta)}(\lambda_j) \right\} \right)^{-1} \zeta \right]^{\frac{1}{2}}} = \frac{\widehat{t}_m(\gamma, \delta, \omega_{1.2})}{(\widehat{\omega}_{1.2}/\omega_{1.2})^{1/2}}. \quad (22)$$

Remark 8. Imposing conditions on m^* such that $\widehat{\omega}_{1.2} \rightarrow_p \omega_{1.2}$, then $\widehat{t}_m(\gamma, \delta, \widehat{\omega}_{1.2}) \rightarrow_d N(0, 1)$. Alternatively, we will keep m^* fixed with the aim of obtaining a better approximation to the sampling distribution of $\widehat{t}_m(\gamma, \delta, \widehat{\omega}_{1.2})$. As in Hualde and Iacone (2017), the main implication is that, when consider a fixed- m limit, $\widehat{\omega}_{1.2}$ loses consistency. In fact, the limit is random, but it is proportional to $\omega_{1.2}$, so, in the limit, the dependence of $\widehat{t}_m(\gamma, \delta, \widehat{\omega}_{1.2})$ on $\omega_{1.2}$ cancels out. However, the main complication is the potential dependence between the limits of $\widehat{t}_m(\gamma, \delta, \omega_{1.2})$ (that is a $N(0, 1)$ random variable) and $\widehat{\omega}_{1.2}$, because, if the limits were dependent, $\widehat{t}_m(\gamma, \delta, \widehat{\omega}_{1.2})$ would not be asymptotically pivotal. Actually, avoiding this dependence is the reason why Vogelsang and Wagner (2014) need to construct new residuals (adjusted residuals in their terminology, see p. 746) by a relatively convoluted method, achieving the desired independence (see their last paragraph in their Lemma 2). Fortunately, this type of adjustment is not necessary in our setting and our very simple estimator (21) satisfies the crucial independence condition, as Theorem 5 shows.

Define $W(r) = (W_1(r), W_2(r))'$ as a Brownian motion with covariance matrix I_2 , $H(r; \beta) = (W_2(r; \beta), W_2(r))'$, where

$$W_2(r; \beta) = \int_0^r \frac{(r-s)^{\beta-1}}{\Gamma(\beta)} dW_2(s),$$

$$\widehat{\Psi}_{\sin}(j) = \int_0^1 \sin(2\pi jr) \widehat{H}(r; \beta + 1) dr, \quad \widehat{\Psi}_{\cos}(j) = \int_0^1 \cos(2\pi jr) \widehat{H}(r; \beta + 1) dr,$$

and

$$\begin{aligned}\Psi_m(\beta) &= H(1; \beta + 1) H'(1; \beta + 1) + 2 \sum_{j=1}^m (2\pi j)^2 \left(\widehat{\Psi}_{\sin}(j) \widehat{\Psi}'_{\sin}(j) + \widehat{\Psi}_{\cos}(j) \widehat{\Psi}'_{\cos}(j) \right), \\ \psi_m(r; \beta) &= H(1; \beta + 1) + 2 \sum_{j=1}^m 2\pi j \left(\widehat{\Psi}_{\sin}(j) \cos(2\pi j r) - \widehat{\Psi}_{\cos}(j) \sin(2\pi j r) \right).\end{aligned}$$

Also, let

$$\Xi_0(\beta) = \left(\int_0^1 dW_1(r) - H'(1; \beta + 1) \Psi_m^{-1}(\beta) \int_0^1 \psi_m(r; \beta) dW_1(r) \right)^2$$

and for $j = 1, \dots, m^*$,

$$\begin{aligned}\Xi_j(\beta) &= \left(\int_0^1 \cos(2\pi j r) dW_1(r) - 2\pi j \widehat{\Psi}'_{\sin}(j) \Psi_m^{-1}(\beta) \int_0^1 \psi_m(r; \beta) dW_1(r) \right)^2 \\ &\quad + \left(\int_0^1 \sin(2\pi j r) dW_1(r) + 2\pi j \widehat{\Psi}'_{\cos}(j) \Psi_m^{-1}(\beta) \int_0^1 \psi_m(r; \beta) dW_1(r) \right)^2\end{aligned}$$

Theorem 5. *Let Assumption 1 hold and m, m^* be any nonnegative fixed numbers such that $m^* \leq m$. Then, as $n \rightarrow \infty$,*

$$\widehat{t}_m(\gamma, \delta, \widehat{\omega}_{1.2}) \rightarrow_d \frac{Z}{\left(\frac{1}{2m^*+1} \sum_{j=0}^{m^*} c_j \Xi_j(\beta) \right)^{1/2}} \quad (23)$$

where $Z \sim N(0, 1)$, and numerator and denominator are independent random variables.

Remark 9. As shown in the proof, the independence between the numerator and denominator of the limit in (23) relies on the conditional uncorrelation between the limit of $n^\beta (\widehat{\nu}_m(\gamma, \delta) - \nu)$ and those of $n^{-1/2} \sum_{t=1}^n \widetilde{u}_{1.2t} \cos(\lambda_j t)$ and $n^{-1/2} \sum_{t=1}^n \widetilde{u}_{1.2t} \sin(\lambda_j t)$ for $j = 0, \dots, m^*$. Due to the orthogonality properties of the trigonometric functions, this only occurs whenever $j \leq m$, which justifies the restriction $m^* \leq m$. As it is evident from (23), the quantiles from the corresponding distribution need to be simulated. In addition, β is typically unknown, so, in practice, a feasible strategy is to compute the critical values for an estimated β . This is precisely the plug-in approach

for obtaining critical values employed by McElroy and Politis (2012, 2013) in a related setting.

Finally, we present below corresponding results for the fully feasible $\hat{\nu}_m(\hat{\gamma}, \hat{\delta})$, $\hat{\rho}_m(\hat{\gamma}, \hat{\delta})$ and $\hat{t}_m(\hat{\gamma}, \hat{\delta}, \hat{\omega}_{1.2})$.

Theorem 6. *Let Assumptions 1, 2 hold and m, m^* be any nonnegative fixed numbers. Then, as $n \rightarrow \infty$, results (20) and (23) apply for $(\hat{\nu}_m(\hat{\gamma}, \hat{\delta}), \hat{\rho}_m(\hat{\gamma}, \hat{\delta}))'$ and $\hat{t}_m(\hat{\gamma}, \hat{\delta}, \hat{\omega}_{1.2})$, respectively.*

Again, this result can be justified as part (ii) of Theorem 4, so its proof is omitted.

5 Monte Carlo evidence

We investigated the properties of the sampling distribution of test statistics $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$ and $\hat{t}_m(\hat{\gamma}, \hat{\delta}, \hat{\omega}_{1.2})$ and their approximation by the limit given in Theorems 5 and 6 by means of a Monte Carlo exercise.

Setting $m = m^*$ throughout, we first illustrate some characteristics of this limit. In particular, we reported in Table 1 the 95%, 97.5%, 99.5% quantiles (denoted as $\alpha = 0.10, 0.05, 0.01$, respectively) of the limit distribution for various values of m , and $\beta = 0.1 + 0.1k$, $k = 0, 1, \dots, 19$, so that β spanned 0.1 to 2. The limit distribution was simulated using Gauss: we simulated the random variable distributed as in the limit (23) by approximating integrals by summations of 10,000 steps, and using standard normally distributed random variates in the stochastic integrals, and we then simulated the limit distribution by looking at the empirical distribution of 10,000 limit variates (for $m = 1$ we used 100,000 variates, because we observed a certain instability in the upper 1% quantile as β was changed). There are two interesting features to be noted. First, the fixed- m limit has thicker tails than the $N(0, 1)$, this being extremely notorious for the $m = 1$ case. Indeed, we found that tails are thicker even when compared to the t_{2m} distribution. However, as expected, the quantiles tend to those of a $N(0, 1)$ as m increases. Second, quantiles are relatively invariant to β , which, given that in most practical situations β will be considered unknown, favours heavily the application of the previously mentioned plug-in approach for obtaining critical values. We illustrate these facts in Figures 1 and 2, respectively: in Figure 1 we compare the cumulative distributions of the limit in (23) for $\beta = 0.4, 1, 2$ to the $N(0, 1)$ (Z) when $m = 1$, and in Figure 2 we set $\beta = 1$ and compare the distributions for $m = 1, 2, 5$ to the standard normal. In particular, it is evident that the limit distributions always shrink towards the $N(0, 1)$ as m increases, that the limit for $m = 1$ has a much bigger dispersion than for $m > 1$, and that the effect of

β is very minor, to the point that three plots in Figure 1 seem to completely overlap.

Next, in Figures 3 to 6 we compare the sampling distribution of $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1,2})$ for $n = 64, 256, 1024$ with the corresponding fixed- m limit in four scenarios: Figure 3, $m = 1, \beta = 0.4$; Figure 4, $m = 1, \beta = 1$; Figure 5, $m = 2, \beta = 1$; Figure 6, $m = 3, \beta = 1$. The sampling cumulative distributions have been computed from 10,000 repetitions of the test statistics $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1,2})$ for the simulated DGP (1), (2) with

$$u_{1t} = \phi_1 u_{1t-1} + \eta_{1t}, \quad u_{2t} = \phi_2 u_{2t-1} + \eta_{2t},$$

where

$$\eta_{1t} = e_{1t}, \quad \eta_{2t} = \frac{\theta}{\sqrt{\theta^2 + (1-\theta)^2}} e_{1t} + \frac{1-\theta}{\sqrt{\theta^2 + (1-\theta)^2}} e_{2t}$$

and $e_t = (e_{1t}, e_{2t})'$ is normally independently distributed with $E(e_t) = 0$, $E(e_t e_t') = I_2$. In particular we used $\gamma = 0$, $\delta = \beta$, $\phi_1 = 0.3$, $\phi_2 = 0.7$, $\theta = 0.5$, and $\nu = 1$. Notice that $E(\eta_{1t}^2) = 1$, $E(\eta_{2t}^2) = \frac{\theta^2}{\theta^2 + (1-\theta)^2} + \frac{(1-\theta)^2}{\theta^2 + (1-\theta)^2} = 1$ and $E(\eta_{1t} \eta_{2t}) = \frac{\theta}{\sqrt{\theta^2 + (1-\theta)^2}}$, so this model can be casted in a way consistent with Assumption 1. With this DGP we imposed a relevant degree of autocorrelation, especially in the explanatory variable, as well as correlation between the error and the regressor in (1), that makes the estimation of ρ very important. Notice also that at this stage we treat γ and δ (and therefore β) as known, thus making the test statistic infeasible (except, perhaps, for the $\gamma = 0$, $\delta = 1$ case): we do so to distinguish the asymptotic properties of the test statistic from the interference due to the estimation of γ and δ , but we will later consider the feasible case of unknown γ and δ .

As expected, the empirical distribution of the test statistic is closer to the limit as n increases, but, nicely, the limit provides an accurate approximation to the sampling distribution of the test statistic even for $n = 64$. Comparing Figure 3 and Figure 4 more in detail, however, we can see that, especially for the $n = 64$ case, the sampling distribution for $\beta = 1$ approximates the limit better than for the $\beta = 0.4$ case, perhaps reflecting the faster rate of convergence of the estimate of ν in the former case. Another noticeable feature arises from the comparison between Figures 4, 5 and 6: for given sample size, the sampling distribution is closest to the corresponding limit distribution when $m = 1$, and most distant when $m = 3$, hinting at the fact that inference based on estimates with larger bandwidth m may be more exposed to size distortion. At least for the values of m considered in these three figures, however, this seems to be a very minor effect.

In Figures 7 to 11 we provide corresponding results for the feasible statistic $\hat{t}_m(\hat{\gamma}, \hat{\delta}, \hat{\omega}_{1,2})$: in particular, for Figures 7 to 10 we consider the same values of β

and m as in Figures 3 to 6, to study the effect of replacing the unknown β with estimates; in Figure 11 we consider the larger bandwidth $m = 10$. Consistently with the semiparametric nature of the test statistic, we computed $\hat{\gamma}, \hat{\delta}$ as the local Whittle estimates (initializing the numerical optimization taking the log-periodogram regression estimate as the starting value), with bandwidth set to $\lfloor n^{0.65} \rfloor$, as recommended for example in Abadir, Distaso and Giraitis (2007). Typically, this bandwidth is much larger than our choices for m , as for the estimates $\hat{\gamma}, \hat{\delta}$ we exploit a consistency argument in Theorem 6. Local Whittle and log-periodogram regression estimation require prior knowledge of whether the series should be used in levels or in first differences (or, for $\delta = 2$, even in second differences): we took as known this piece of information, but this is a mild assumption. Notice that while δ can be computed directly from x_t , series $\Delta^{-\gamma} \{u_{1t} 1(t > 0)\}$ is not observable: we estimated u_{1t} from residuals after estimating ν by NBLs, with bandwidth set as m . Overall results are worse than in the infeasible case, in the sense that, for the same β and m and sample size, the distribution of the feasible statistic is more distant from the limit than the distribution of the infeasible one. However, again, results improve as n increases. Also, in comparison to the infeasible situation, the improvement as β increases is much more evident. This reflects the fact that the first stage estimate of ν (NBLs), which is employed in the estimation of γ , has a better rate of convergence the larger is β .

Finally in Figure 12, we summarize a small investigation of local power: here we focussed again on the infeasible $\beta = 1$ case, but we modified the DGP using $\nu = (1 + c/n)$ for c taking values between 0 and 100, $n = 512$, and we tested the incorrect null hypothesis $H_0 : \nu = 1$. For this exercise only, we set $\phi_1 = 0$, $\phi_2 = 0$ and $\theta = 0$ in the DGP. We considered bandwidths $m = 1$, $m = 2$, $m = 3$ and $m = n/2$, the latter mimicking the properties of OLS estimation. Notice that OLS is in general inconsistent, but with this particular assumption on ϕ_1 and θ it is actually efficient, so the local power provides a benchmark, against which we can measure the performance of bandwidths such as $m = 1$, $m = 2$ and $m = 3$. Consistently with findings in the literature, the power in our exercise is increasing in m , but the power loss when $m = 2$ or $m = 3$ is very small, which is certainly an encouraging result.

Having thus established that the limit distribution from Theorems 5 and 6 are broadly a suitable approximation and a marked improvement on the standard normal, and that this is also true for small bandwidths, including $m=1$, in the second part of the Monte Carlo exercise we investigate more in detail the empirical size.

In Tables 2 to 5 we report the empirical sizes corresponding to nominal $\alpha = 0.05$,

$\alpha = 0.10$, for $m = 1, 2, 3, 5, 10, \lfloor \sqrt{n} \rfloor$, noting that for this last choice of m our theory does not apply, and $n = 64, 128, 256, 512$, for the (correct) null hypothesis $H_0 : \nu = 1$ vs. alternative $H_A : \nu \neq 1$. Notice that for some cases with $m = \lfloor \sqrt{n} \rfloor$ the critical values are not reported in Table 1 (e.g., $\lfloor \sqrt{512} \rfloor$ requires $m = 22$): we excluded these from Table 1 for the benefit of the presentation but the appropriate critical values were used even in those cases. We considered the same correlated DGP as in the previous study, thus setting $\phi_1 = 0.3$, $\phi_2 = 0.7$, $\theta = 0.5$; however, we expanded the range of values of γ and δ , and considered $(\gamma, \delta) = (0.2, 0.4), (0, 0.4), (0.4, 1.2), (0, 1), (0.4, 2)$. We therefore cover both cases with weak cointegration ($\beta < 1/2$) and strong cointegration (with $\beta > 1/2$), including the familiar $\gamma = 0, \delta = 1$ case: the wide range of values for β is interesting, as we conjecture that the empirical size will approximate the theoretical 5% level better for larger values of β , because the faster rate of convergence of the first stage NBLS estimate of ν will result in less contamination of the residuals from the correlation with the explanatory variable; the two specifications when $\beta < 1/2$ are also interesting as the $\gamma = 0, \delta = 0.4$ case meets the condition $\gamma + \delta < 1/2$ and will therefore allow a comparison with the test statistic in Christensen and Nielsen (2006). Other combinations of (γ, δ) with the same β as in the cases we include in the experiment led to almost identical results.

In each table we present the empirical sizes when critical values are obtained from the fixed- m limit (denoted as F in the tables) or $N(0, 1)$, for both the infeasible statistic $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$ and for the feasible statistic $\hat{t}_m(\hat{\gamma}, \hat{\delta}, \hat{\omega}_{1.2})$. We present results for the 5% size in Tables 2 and 3, and for the 10% size in Tables 4 and 5, the infeasible statistics being covered in Tables 2 and 4, and the feasible ones in Tables 3 and 5.

The size performance for the bandwidth $m = 1$ is impressive: for the infeasible statistic $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$ the empirical sizes are extremely close to the nominal ones even in the smallest sample. Remarkably, although results slightly worsen, an almost identical pattern is obtained for the feasible $\hat{t}_m(\hat{\gamma}, \hat{\delta}, \hat{\omega}_{1.2})$.

As it was anticipated, results deteriorate as m gets larger, especially when the sample is smaller, when the cointegration gap β is small and when the feasible statistic is considered. Thus, for example, the size properties are still very good when $m = 2$ and the infeasible statistic is considered, with only a very minor distortion for $\beta=0.2$ or $\beta=0.4$ and $n=64$, whereas for $m = 10$ results are more subject to size distortion, especially in smaller samples or for small β . However, even for $m = 10$ we find that for large enough n the size properties are good or acceptable, at least in the case of the infeasible statistic, thus confirming that the limit in Theorem 5 provides a very valuable guidance. On the other hand, inference based on the critical values from the

$N(0, 1)$ is heavily distorted. For obvious reasons, this effect is worse when $m=1$, and it is somehow mitigated as m increases, but the fixed- m inference is superior even for the largest bandwidth.

The last case we consider is the study of the size when the bandwidth rule $m = \lfloor \sqrt{n} \rfloor$ is adopted. This case is particularly of interest as it seems closer to the spirit of the bandwidth choice in Hualde and Robinson (2006): here too we find that the critical values assuming the fixed- m limit provide better size than assuming standard asymptotics, although in this case we still find size distortion even in the largest samples. This finding supports the widespread conclusion that fixed smoothing asymptotics systematically help improving size when compared to standard asymptotics.

Overall, the outcome of this Monte Carlo exercise is consistent with the two main results that are commonly found in similar studies, namely, that the critical values from the simulation of Theorem 5 help correcting size for any bandwidth choice, and that there is a trade-off between size and power, as the best size is obtained choosing small bandwidths, but the best power would require larger bandwidths (as it is evident from Figure 12). Also, as m increases, the large- m limiting theory is favoured and this is clearly reflected in our results. However, using fixed- m critical values always lead to a better size performance. Both approaches (large- m and fixed- m) seem to be equally affected by the distortion due to estimating γ , δ , although, in view of (16), this effect was clearly expected just for the large- m theory, especially for small β .

We complement our Monte Carlo exercise with a comparison with another semi-parametric method, that allows for inference on ν using the NBLS estimate $\hat{\nu}_m$. Under the additional assumption that Ω is block diagonal, Christensen and Nielsen (2006) established that, when $\gamma + \delta < 1/2$,

$$\sqrt{m}\lambda_m^{\gamma-\delta}(\hat{\nu}_m - \nu) \rightarrow_d N\left(0, \frac{\omega_{11}(1-2\delta)^2}{2\omega_{22}(1-2\delta-2\gamma)}\right),$$

which naturally leads to the test statistic

$$\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22}) = \sqrt{m}\lambda_m^{\gamma-\delta} \left(\frac{\hat{\omega}_{11}(1-2\delta)^2}{2\hat{\omega}_{22}(1-2\delta-2\gamma)} \right)^{-1/2} (\hat{\nu}_m - \nu),$$

for appropriate estimates $\hat{\omega}_{11}$, $\hat{\omega}_{22}$, of ω_{11} , ω_{22} , respectively. In comparison to the inferential approach based on $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1,2})$, using $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$ may be appealing as it is based on the simpler estimator $\hat{\nu}_m$, but it has the cost of exposing the inference to the risk of having in fact non-zero ρ .

We then compare the performances of the two statistics by looking again at size and including now more evidence about power behaviour, for the case $\gamma = 0$, $\delta = 0.4$. Parameters ω_{11} and ω_{22} are estimated averaging at low frequencies the periodograms of the NBLS residuals and of $\Delta^\delta x_t$, respectively, with bandwidth set at $[n^{0.65}]$. The inference is still infeasible in the sense that we do not estimate γ and δ and consider $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$ and $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$ instead, but this avoids the problem of dealing with estimates of $\delta + \gamma$ exceeding $1/2$, which would be problematic for the statistic $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$. Christensen and Nielsen (2006) do not give a bandwidth rule, but in their empirical example they use bandwidths spanning $m = 3$ to $m = 15$ for $n = 417$ and in general recommend to choose low values for m , see page 360: with a bandwidth rule $m = [n^a]$, the choice $m = 15$ corresponds to $a = \ln(15)/\ln(417) \approx 0.45$ so, considering also the advice to lean towards low values for m , we take $m = [n^{0.4}]$. For the test based on the $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$ we take $m=3$ and $m=10$: the former choice should have better size properties when Ω is not block diagonal, whereas the latter could have better power.

For the size study, we take samples of $n=64$, 128, 256 and 512, and we consider three DGPs, denoted as DGP I, DGP II and DGP III: for DGP I we assume no dependence and no correlation between x_t and u_{1t} , thus setting $\phi_1 = 0$, $\phi_2 = 0$, and $\theta = 0$; for DGP II we imposed dependence but no correlation, thus setting $\phi_1 = 0.3$, $\phi_2 = 0.7$, $\theta = 0$, and for DGP III we also restored correlation by introducing $\theta = 0.5$. For each combination of DGP's and sample size we repeat the experiment 10,000 times. Results for the size study are reported in Table 6, where the empirical size for $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$ is summarized in columns with heading CN, and that for $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$ (with fixed- m asymptotics) appears in columns with heading F. The most noticeable finding from Table 6 is that, as expected, the size for the $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$ statistic is heavily distorted in presence of endogeneity, the empirical size spuriously increasing towards 1 as the sample gets bigger: Christensen and Nielsen (2006) required that Ω is block diagonal, and Nielsen and Frederiksen (2011) showed that, otherwise, the NBLS estimate $\hat{\nu}_m$ is subject to a lower order bias, which is therefore responsible for this spurious significance. More worrying still, perhaps, is the fact that, even when endogeneity is not a problem (as it is the case for DGP I and DGP II), the test based on the $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$ statistic is still subject to size distortion: admittedly, this is reducing as the sample size is increased, but it remains relevant even in the largest sample ($n=512$). On the other hand, the presence of weak autocorrelation in the explanatory variables or in the unobserved component does not have a strong adverse effect on size. In comparison, fixed- m critical values coupled

with the $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$ statistic always yield better (or even much better) empirical size, the only evidence of some size distortion being when the largest bandwidth ($m=10$) and some endogeneity affects the data, as in DGP III. Even then, however, the size is broadly correct in the largest sample.

For the power study, we consider DGP I, testing as before $H_0 : \nu = 1$, when in fact in the DGP we used $\nu = 1 + c/n$ for various values of c and $n=512$. In view of the size distortion of $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$, we used the size-adjusted power for the test using this statistic. Also, to facilitate comparison between $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$ and $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$, in this part of the experiment we used the same bandwidth, and we set $m=10$. Results are given in Table 7, where we also display the power for a third statistic,

$$\tilde{\tau}_n = \left(\sum_{j=0}^{n/2} c_j I_{xx}(\lambda_j) \right)^{1/2} (\hat{\omega}_{11})^{-1} (\hat{\nu}_n - 1)$$

which is the standardized OLS estimate. Unlike with $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$, the power results for $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$ or $\tilde{\tau}_n$ were not size-adjusted: in both cases the 5% empirical size was approximately correct (for $\tilde{\tau}_n$ it was 0.0501), so there was no need of size-adjusting. Not surprisingly, given that $\rho = 0$ is imposed, the statistic based on OLS displays the highest power. Also, $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$ displays more power than $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$, so, given that both test statistics use the same bandwidth and considering that $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$ is size corrected, the superior power performance of the this test is not spurious and it is not due to a different bandwidth choice.

6 Conclusions

The main contribution of the paper is to present three limiting results involving fractionally cointegrated processes where the bandwidth, unlike in the traditional limiting theory, is kept fixed. First we derived the fixed- m limit of the NBLS estimator. As with the more traditional large- m theory, this limit is, in general, not useful to carry out standard inference. Thus we next derive the “partial” fixed- m limiting theory for a full system GLS-type version of the NBLS estimator which removes the endogeneity bias. This theory is called “partial” in the sense that it relies on consistency arguments for the estimators of nuisance parameters.

Finally, we discussed the more “complete” fixed- m limiting theory for that estimator, in the sense that fixed- m arguments were used for the estimators of the short run component of the model $(\rho, \omega_{1.2})$, therefore without relying on consistency arguments

and taking into account the inherent randomness of the estimators of $(\rho, \omega_{1,2})$ in the asymptotic distribution. Note however that our “complete” theory relies on consistency of the estimators of the integration orders, which seems unavoidable. Nicely, our Monte Carlo evidence suggests that the fixed- m limiting theory leads to a much more accurate approximation to the sampling distribution of the discussed estimators and test statistics.

APPENDIX

Proof of Theorem 1

The proof follows from almost identical arguments to those in the proof of Lemma 1 of Hualde and Iacone (2017) with the main difference that now the fractional invariance principle of Hosoya (2005) needs to be employed. This just implies that the Type 2 (instead of the Type 1) fractional Brownian motion characterizes the limit in (6).

Proof of Corollary 1

Given the mixed normality of the limit in (12), the result follows by showing that the conditional variance of $N_m(\Omega, \beta)$ (given $S(r; \beta + 1)$) is $D_m(\Omega, \beta)$. For $m = 0$ the proof is immediate. Next, for $m > 0$, first note that for any $j = 1, \dots, m$,

$$\int_0^1 r \cos(2\pi jr) dr = 0, \quad \int_0^1 r \sin(2\pi jr) dr = -(2\pi j)^{-1},$$

so

$$\int_0^1 \sin(2\pi js) A(1) \widehat{B}(s) ds = \int_0^1 \sin(2\pi js) A(1) B(s) ds + \frac{1}{2\pi j} A(1) B(1), \quad (24)$$

$$\int_0^1 \cos(2\pi js) A(1) \widehat{B}(s) ds = \int_0^1 \cos(2\pi js) A(1) B(s) ds. \quad (25)$$

Next, by integration by parts

$$\begin{aligned} \int_0^1 \sin(2\pi js) A(1) B(s) ds &= -A(1) \int_0^1 \frac{1}{2\pi j} (1 - \cos(2\pi js)) dB(s) \\ &= -\frac{1}{2\pi j} A(1) B(1) + A(1) \int_0^1 \frac{1}{2\pi j} \cos(2\pi js) dB(s), \end{aligned} \quad (26)$$

so noting (24)

$$\int_0^1 \sin(2\pi js) A(1) \widehat{B}(s) ds = A(1) \int_0^1 \frac{1}{2\pi j} \cos(2\pi js) dB(s).$$

Similarly, noting (25) and using integration by parts

$$\int_0^1 \cos(2\pi js) A(1) \widehat{B}(s) ds = -A(1) \int_0^1 \frac{1}{2\pi j} \sin(2\pi js) dB(s). \quad (27)$$

Thus

$$\begin{aligned} N_m(\Omega, \beta) = & \zeta' \Omega^{-1} \left\{ S(1; \beta + 1) A(1) \int_0^1 dB(r) + 2 \sum_{j=1}^m 2\pi j \right. \\ & \times \left[\int_0^1 \sin(2\pi jr) \widehat{S}(r; \beta + 1) dr A(1) \int_0^1 \cos(2\pi js) dB(s) \right. \\ & \left. \left. - \int_0^1 \cos(2\pi jr) \widehat{S}(r; \beta + 1) dr A(1) \int_0^1 \sin(2\pi js) dB(s) \right] \right\}. \end{aligned}$$

Then noting that for any integer $j, k \geq 1$,

$$\int_0^1 \cos(2\pi jr) dr = \int_0^1 \sin(2\pi jr) dr = \int_0^1 \cos(2\pi jr) \sin(2\pi kr) dr = 0, \quad (28)$$

whereas for any $j \neq k$,

$$\int_0^1 \cos(2\pi jr) \cos(2\pi kr) dr = \int_0^1 \sin(2\pi jr) \sin(2\pi kr) dr = 0, \quad (29)$$

the conditional variance of $N_m(\Omega, \beta)$ (given $S(r; \beta + 1)$) is

$$\begin{aligned} & \zeta' \Omega^{-1} \zeta \left\{ S^2(1; \beta + 1) + 4 \sum_{j=1}^m (2\pi j)^2 \left(\int_0^1 \sin(2\pi jr) \widehat{S}(r; \beta + 1) dr \right)^2 \int_0^1 \cos^2(2\pi jr) dr \right. \\ & \left. + 4 \sum_{j=1}^m (2\pi j)^2 \left(\int_0^1 \cos(2\pi jr) \widehat{S}(r; \beta + 1) dr \right)^2 \int_0^1 \sin^2(2\pi jr) dr \right\}, \end{aligned}$$

which equals $D_m(\Omega, \beta)$ because for any $j = 1, \dots, m$,

$$\int_0^1 \cos^2(2\pi jr) dr = \int_0^1 \sin^2(2\pi jr) dr = \frac{1}{2}. \quad (30)$$

Proof of Theorem 3

Part (i) can be justified as in Hualde and Robinson (2010), noting that their arguments also hold under our slightly milder Assumption 1. Next, noting that for $i = 1, 2$, $\sum_{s=1}^t u_{is}(c) = u_{it}(c-1)$, part (ii) follows by our Theorem 2 and Lemma C.5 of Robinson and Hualde (2003).

Proof of Theorem 4

Noting that

$$\begin{pmatrix} \widehat{\nu}_m(\gamma, \delta) - \nu \\ \widehat{\rho}_m(\gamma, \delta) - \rho \end{pmatrix} = \left(\operatorname{Re} \left\{ \sum_{j=0}^m c_j I_{r(\gamma, \delta)}(\lambda_j) \right\} \right)^{-1} \operatorname{Re} \left\{ \sum_{j=0}^m c_j I_{r(\gamma, \delta)u_{1.2}}(\lambda_j) \right\},$$

(20) follows from identical arguments as those in the proof of Theorem 2 applying summation by parts twice as in (26), (27).

Proof of Corollary 3

The result follows immediately by showing that

$$\int_0^1 \phi_m(r; \Omega, \beta) \phi'_m(r; \Omega, \beta) dr = \Phi_m(\Omega, \beta), \quad (31)$$

but this can be easily justified using results (28)-(30).

Proof of Theorem 5

Considering the second equality in (22), the convergence of the numerator of $\widehat{t}_m(\gamma, \delta, \omega_{1.2})$ to Z follows by Corollary 3. Next,

$$\widehat{\omega}_{1.2} = \frac{1}{2m^* + 1} \sum_{j=0}^{m^*} c_j \left[\left(\frac{1}{n^{1/2}} \sum_{t=1}^n \widetilde{u}_{1.2t} \cos(\lambda_j t) \right)^2 + \left(\frac{1}{n^{1/2}} \sum_{t=1}^n \widetilde{u}_{1.2t} \sin(\lambda_j t) \right)^2 \right],$$

where

$$\widetilde{u}_{1.2t} = u_{1.2t} - (\widehat{\nu}_m(\gamma, \delta) - \nu, \widehat{\rho}_m(\gamma, \delta) - \rho) r_t(\gamma, \delta).$$

By Theorem 3 and using identical arguments as in the proof of Theorem 2 and Corollary 1, it is straightforward to show that for any $j = 1, \dots, m^*$,

$$\frac{1}{n^{1/2}} \sum_{t=1}^n \widetilde{u}_{1.2t} \rightarrow_d l^{(1)}(\Omega, \beta), \quad (32)$$

$$\frac{1}{n^{1/2}} \sum_{t=1}^n \widetilde{u}_{1.2t} \cos(\lambda_j t) \rightarrow_d l_j^{(2)}(\Omega, \beta), \quad (33)$$

$$\frac{1}{n^{1/2}} \sum_{t=1}^n \tilde{u}_{1.2t} \sin(\lambda_j t) \rightarrow_d l_j^{(3)}(\Omega, \beta), \quad (34)$$

where

$$\begin{aligned} l^{(1)}(\Omega, \beta) &= \int_0^1 dE(r) - F'(1; \beta + 1) \Phi_m^{-1}(\Omega, \beta) \int_0^1 \phi_m(r; \Omega, \beta) dE(r), \\ l_j^{(2)}(\Omega, \beta) &= - \int_0^1 \cos(2\pi jr) dE(r) + 2\pi j \hat{\Phi}'_{\sin}(j) \Phi_m^{-1}(\Omega, \beta) \int_0^1 \phi_m(r; \Omega, \beta) dE(r), \\ l_j^{(3)}(\Omega, \beta) &= - \int_0^1 \sin(2\pi jr) dE(r) - 2\pi j \hat{\Phi}'_{\cos}(j) \Phi_m^{-1}(\Omega, \beta) \int_0^1 \phi_m(r; \Omega, \beta) dE(r). \end{aligned}$$

Then, noting that $E(r)$ and $F(r; \beta + 1)$ are independent processes, by the continuous mapping theorem

$$\hat{\omega}_{1.2} \rightarrow_d \frac{\omega_{1.2}}{2m^* + 1} \sum_{j=0}^{m^*} c_j \Xi_j(\beta). \quad (35)$$

Finally we justify the independence between the numerator and denominator of (23). Clearly the limit in (35) is a functional of the limits in (32)-(34). We first show that the conditional variance (given $F(r; \beta + 1)$), say $Cov_F(\cdot, \cdot)$, between the limit of $n^\beta(\hat{\nu}_m(\gamma, \delta) - \nu)$ and each of the limits in (32)-(34) is zero for $j \leq m$.

First, using (28) and (31),

$$\begin{aligned} & Cov_F \left(l^{(1)}(\Omega, \beta), \zeta' \Phi_m^{-1}(\Omega, \beta) \int_0^1 \phi_m(r; \Omega, \beta) dE(r) \right) \\ &= \omega_{1.2} \left(\int_0^1 \phi'_m(r; \Omega, \beta) dr - F'(1; \beta + 1) \right) \Phi_m^{-1}(\Omega, \beta) \zeta = 0. \end{aligned}$$

Next,

$$\begin{aligned}
& Cov_F \left(l_j^{(2)}(\Omega, \beta), \zeta' \Phi_m^{-1}(\Omega, \beta) \int_0^1 \phi_m(r; \Omega, \beta) dE(r) \right) \\
&= \omega_{1.2} \left(- \int_0^1 \phi'_m(r; \Omega, \beta) \cos(2\pi jr) dr + 2\pi j \widehat{\Phi}'_{\sin}(j) \right) \Phi_m^{-1}(\Omega, \beta) \zeta \\
&= \omega_{1.2} \left(-2 \sum_{k=1}^m 2\pi k \int_0^1 \left(\widehat{\Phi}'_{\sin}(k) \cos(2\pi kr) - \widehat{\Phi}'_{\cos}(k) \sin(2\pi kr) \right) \cos(2\pi jr) dr \right. \\
&\quad \left. + 2\pi j \widehat{\Phi}'_{\sin}(j) \right) \Phi_m^{-1}(\Omega, \beta) \zeta \\
&= \omega_{1.2} \left(-2\pi j \widehat{\Phi}'_{\sin}(j) + 2\pi j \widehat{\Phi}'_{\sin}(j) \right) \Phi_m^{-1}(\Omega, \beta) \zeta = 0,
\end{aligned} \tag{36}$$

by (28)-(30), where the two last equalities in (36) are true just if $j \leq m$. Similarly

$$\begin{aligned}
& Cov_F \left(l_j^{(3)}(\Omega, \beta), \zeta' \Phi_m^{-1}(\Omega, \beta) \int_0^1 \phi_m(r; \Omega, \beta) dE(r) \right) \\
&= \omega_{1.2} \left(- \int_0^1 \phi'_m(r; \Omega, \beta) \sin(2\pi jr) dr - 2\pi j \widehat{\Phi}'_{\cos}(j) \right) \Phi_m^{-1}(\Omega, \beta) \zeta = 0,
\end{aligned}$$

as in the proof of (36).

Then, by Gaussianity, the conditional uncorrelation becomes conditional independence and given that Z does not depend on $F(r; \beta + 1)$, the limits of $\widehat{t}_m(\gamma, \delta, \omega_{1.2})$ and $\widehat{\omega}_{1.2}$ are unconditionally independent, to conclude the proof of the theorem.

REFERENCES

- Abadir, K., M., Distaso, W., and L. Giraitis, 2007. Nonstationarity-extended local Whittle estimation, *Journal of Econometrics* 141, 1353-1138.
- Andersen, T. G., and R.T. Varneskov, 2018. Consistent inference for predictive regressions in persistent VAR economies. Unpublished manuscript, Northwestern University.
- Bester, C.A., Conley, T.G., and C.B. Hansen, 2011. Inference with dependent data using cluster covariance estimators, *Journal of Econometrics* 165, 137-151.
- Bunzel, H., 2006. Fixed- b asymptotics in single-equation cointegration models with endogenous regressors, *Econometric Theory* 22, 743-755.

- Bunzel, H., and T.J. Vogelsang, 2005. Powerful trend function tests that are robust to strong serial correlation, with an application to the Prebisch-Singer hypothesis, *Journal of Business and Economic Statistics* 23, 381-394.
- Chen, W.W., and C.M. Hurvich, 2003. Estimating fractional cointegration in the presence of polynomial trends, *Journal of Econometrics* 117, 95-121.
- Christensen, B. J., and M. O. Nielsen, 2006. Asymptotic normality of narrow-band least squares in the stationary fractional cointegration model and volatility forecasting, *Journal of Econometrics* 133, 343-371.
- Hannan, E.J., 1973. The asymptotic theory of linear time-series models *Journal of Applied Probability* 10,130-145.
- Hosoya, Y., 2005. Fractional invariance principle, *Journal of Time Series Analysis* 26, 463-486.
- Hualde, J., and F. Iacone, 2015. Small- b and fixed- b asymptotics for weighted covariance estimation in fractional cointegration, *Journal of Time Series Analysis* 36, 528-540.
- Hualde, J., and F. Iacone, 2017. Fixed bandwidth asymptotics for the studentized mean of fractionally integrated processes, *Economics Letters* 150, 39-43.
- Hualde, J., and P.M. Robinson, 2006. Semiparametric estimation of fractional cointegration. STICERD - Econometrics Paper Series No. EM/2006/502.
- Hualde, J., and P.M. Robinson, 2007. Root- n -consistent estimation of weak fractional cointegration. *Journal of Econometrics* 140, 450-484.
- Hualde, J., and P.M. Robinson, 2010. Semiparametric inference in multivariate fractionally cointegrated systems, *Journal of Econometrics* 157, 492-511.
- Iacone, F., Leybourne, S.J., and A.M. R. Taylor, 2013. On the behaviour of fixed- b trend break tests under fractional integration, *Econometric Theory* 29, 393-418.
- Ibragimov, R., and U.K. Müller, 2010. t -statistic based correlation and heterogeneity robust inference, *Journal of Business and Economic Statistics* 28, 453-468.
- Jakubowski, A., Mémin, J., and G. Pages, 1989. Convergence en loi des suites d'intégrales stochastique sur l'espace \mathbb{D}^1 de Skorohod. *Probabability Theory and Related Fields* 81, 111-137.

- Jin, S., Phillips, P.C.B., and Y. Sun, 2006. A new approach to robust inference in cointegration, *Economics Letters* 91, 300-306.
- Jansson, M., 2004. The error in rejection probability of simple autocorrelation robust tests, *Econometrica* 72, 937-946.
- Johansen, S., 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12, 231-254.
- Johansen, S., 1991. Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models, *Econometrica* 59 1551-1580.
- Johansen, S., and M.O. Nielsen, 2012. Likelihood inference for a fractionally cointegrated vector autoregressive model, *Econometrica* 80, 2667-2732.
- Kiefer, N.M., and T.J. Vogelsang, 2002. Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size, *Econometric Theory* 18, 1350-1366.
- Kiefer, N.M., and T.J. Vogelsang, 2005. A new asymptotic theory for heteroskedasticity - autocorrelation robust tests. *Econometric Theory* 21, 1130-64.
- Kurtz, T. G., and P. Protter, 1991. Weak limit theorems for stochastic integrals and stochastic differential equations. *Annals of Probability* 19, 1035-1070.
- McElroy, T., and D.N. Politis, 2012. Fixed- b asymptotics for the studentized mean from time series with short, long or negative memory. *Econometric Theory* 28, 471-481.
- McElroy, T., and D.N. Politis, 2013. Distribution theory for the studentized mean for long, short, and negative memory time series. *Journal of Econometrics* 177, 60-74.
- Marinucci, D., 2000. Spectral regression for cointegrated time series with long-memory innovations, *Journal of Time Series Analysis* 21, 685-705.
- Marinucci, D., and P.M. Robinson, 2000. Weak convergence of multivariate fractional processes. *Stochastic Processes and their Applications* 86, 103-120.
- Neave, H.R., 1970. An improved formula for the asymptotic variance of spectrum estimates, *Annals of Mathematical Statistics* 41, 70-77.

- Nielsen, M. O., 2005. Semiparametric estimation in time-series regression with long-range dependence, *Journal of Time Series Analysis* 26, 279-304.
- Nielsen, M. O., and P. Frederiksen, 2011. Fully modified narrow-band least squares estimation of weak fractional cointegration, *The Econometrics Journal* 14, 77-120.
- Phillips, P.C.B., 1991a. Spectral regression for cointegrated time series. In W.A. Barnett, J. Powell & G. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*. Cambridge University Press.
- Phillips, P.C.B., 1991b. Optimal inference in cointegrated systems. *Econometrica* 59, 283-306.
- Robinson, P.M., and J. Hualde, 2003. Cointegration in fractional systems with unknown integration orders, *Econometrica* 71, 1727-1766.
- Robinson, P.M., and F. Iacone, 2005. Cointegration in fractional systems with deterministic trends, *Journal of Econometrics* 129, 263-298.
- Robinson, P.M., and D. Marinucci, 2001. Narrow band analysis of nonstationary processes, *Annals of Statistics* 29, 947-86.
- Robinson, P.M., and D. Marinucci, 2003, Semiparametric frequency domain analysis of fractional cointegration. In P.M. Robinson (ed.), *Time Series with Long Memory*. Oxford University Press, Oxford, pp. 334-373.
- Sun, Y., 2013. A heteroskedasticity and autocorrelation robust F test using an orthonormal series variance estimator, *Econometrics Journal* 16, 1-26.
- Sun, Y., 2014. Let's fix it: Fixed- b asymptotics versus small- b asymptotics in heteroskedasticity and autocorrelation robust inference, *Journal of Econometrics* 178, 659-677.
- Sun, Y., Phillips, P.C.B., and S. Jin, 2008. Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing, *Econometrica* 76, 175-194.
- Vogelsang, T.J., and M. Wagner, 2014. Integrated modified OLS estimation and fixed- b inference for cointegrating regressions, *Journal of Econometrics* 178, 741-760.
- Zhang, X., and X. Shao, 2013. Fixed-smoothing asymptotics for time series, *Annals of Statistics* 41, 1329-1349.

TABLE 1. CRITICAL VALUES FOR $m = m^*$

β	$\alpha \backslash m$	1	2	3	4	5	6	8	10	15	20	30	40
0.1	0.10	11.16	3.03	2.38	2.14	2.02	1.96	1.86	1.82	1.76	1.73	1.69	1.68
	0.05	22.28	4.10	3.06	2.69	2.52	2.39	2.27	2.20	2.12	2.07	2.04	2.02
	0.01	109.3	7.48	4.92	4.07	3.71	3.50	3.21	3.04	2.86	2.79	2.71	2.67
0.2	0.10	11.17	3.04	2.37	2.14	2.02	1.95	1.87	1.82	1.76	1.73	1.69	1.68
	0.05	22.37	4.09	3.07	2.69	2.52	2.39	2.28	2.21	2.12	2.07	2.04	2.02
	0.01	112.0	7.61	4.89	4.10	3.71	3.49	3.23	3.05	2.85	2.79	2.71	2.67
0.3	0.10	11.17	3.03	2.37	2.13	2.02	1.95	1.87	1.82	1.76	1.73	1.69	1.68
	0.05	22.25	4.06	3.06	2.69	2.51	2.39	2.28	2.20	2.12	2.07	2.04	2.02
	0.01	112.1	7.59	4.86	4.13	3.71	3.49	3.23	3.05	2.86	2.79	2.71	2.68
0.4	0.10	11.13	3.04	2.37	2.13	2.02	1.95	1.87	1.82	1.76	1.73	1.69	1.68
	0.05	22.44	4.05	3.07	2.70	2.51	2.39	2.28	2.20	2.12	2.07	2.04	2.02
	0.01	113.4	7.73	4.84	4.14	3.70	3.49	3.25	3.05	2.86	2.80	2.71	2.68
0.5	0.10	11.05	3.05	2.37	2.14	2.02	1.95	1.86	1.81	1.76	1.73	1.69	1.68
	0.05	22.44	4.05	3.07	2.69	2.51	2.40	2.28	2.20	2.12	2.08	2.04	2.02
	0.01	113.3	7.65	4.81	4.13	3.68	3.50	3.25	3.06	2.86	2.80	2.71	2.68
0.6	0.10	11.10	3.05	2.36	2.14	2.02	1.95	1.86	1.81	1.76	1.73	1.69	1.68
	0.05	22.51	4.05	3.07	2.70	2.51	2.41	2.28	2.20	2.13	2.07	2.04	2.02
	0.01	110.5	7.59	4.80	4.12	3.66	3.51	3.25	3.07	2.86	2.80	2.71	2.68
0.7	0.10	11.24	3.04	2.36	2.14	2.02	1.95	1.86	1.81	1.76	1.73	1.69	1.67
	0.05	22.78	4.03	3.06	2.70	2.51	2.40	2.27	2.21	2.12	2.07	2.03	2.02
	0.01	114.2	7.51	4.80	4.07	3.64	3.50	3.25	3.06	2.86	2.79	2.72	2.67
0.8	0.10	11.16	3.05	2.36	2.13	2.02	1.95	1.87	1.81	1.76	1.73	1.69	1.67
	0.05	22.59	4.04	3.07	2.71	2.52	2.41	2.28	2.21	2.12	2.07	2.03	2.02
	0.01	111.9	7.39	4.81	4.10	3.66	3.50	3.24	3.06	2.87	2.80	2.72	2.66
0.9	0.10	11.26	3.05	2.36	2.13	2.02	1.95	1.87	1.81	1.75	1.73	1.69	1.67
	0.05	22.65	4.05	3.05	2.71	2.52	2.41	2.27	2.20	2.12	2.07	2.03	2.02
	0.01	112.6	7.57	4.82	4.07	3.65	3.48	3.26	3.06	2.87	2.80	2.72	2.67
1	0.10	11.26	3.04	2.35	2.13	2.02	1.95	1.87	1.82	1.76	1.73	1.69	1.67
	0.05	22.42	4.04	3.06	2.71	2.52	2.41	2.27	2.20	2.12	2.07	2.03	2.02
	0.01	116.2	7.57	4.83	4.06	3.65	3.47	3.25	3.06	2.86	2.80	2.72	2.67
1.1	0.10	11.19	3.04	2.35	2.13	2.02	1.95	1.87	1.82	1.76	1.73	1.69	1.67
	0.05	22.02	4.05	3.05	2.70	2.53	2.41	2.27	2.20	2.12	2.06	2.04	2.02
	0.01	112.2	7.52	4.85	4.11	3.66	3.47	3.26	3.05	2.86	2.80	2.72	2.67
1.2	0.10	11.11	3.04	2.36	2.12	2.02	1.95	1.87	1.82	1.75	1.73	1.69	1.67
	0.05	22.45	4.09	3.04	2.71	2.52	2.41	2.27	2.20	2.12	2.07	2.04	2.02
	0.01	111.4	7.69	4.95	4.12	3.67	3.46	3.26	3.07	2.86	2.80	2.72	2.66
1.3	0.10	11.12	3.03	2.36	2.13	2.02	1.95	1.87	1.83	1.75	1.73	1.69	1.67
	0.05	22.22	4.10	3.04	2.71	2.53	2.42	2.27	2.20	2.12	2.07	2.04	2.02
	0.01	112.9	7.71	4.97	4.13	3.70	3.47	3.26	3.06	2.86	2.80	2.72	2.67
1.4	0.10	11.12	3.02	2.36	2.13	2.02	1.95	1.87	1.82	1.75	1.73	1.70	1.67
	0.05	22.12	4.07	3.04	2.71	2.53	2.41	2.28	2.21	2.11	2.07	2.04	2.02
	0.01	113.4	7.78	5.01	4.12	3.67	3.47	3.24	3.06	2.86	2.80	2.71	2.67
1.5	0.10	11.00	3.02	2.37	2.13	2.02	1.95	1.87	1.82	1.75	1.73	1.69	1.67
	0.05	21.81	4.04	3.05	2.71	2.53	2.42	2.27	2.21	2.12	2.07	2.04	2.02
	0.01	110.1	7.74	5.09	4.15	3.68	3.46	3.24	3.05	2.86	2.80	2.71	2.67
1.6	0.10	11.00	3.01	2.37	2.14	2.02	1.95	1.87	1.82	1.75	1.73	1.69	1.67
	0.05	22.10	4.03	3.06	2.71	2.53	2.41	2.27	2.21	2.12	2.07	2.04	2.02
	0.01	112.4	7.82	5.09	4.13	3.69	3.45	3.23	3.05	2.86	2.81	2.71	2.66
1.7	0.10	11.01	3.00	2.37	2.14	2.02	1.95	1.87	1.82	1.75	1.73	1.69	1.67
	0.05	22.26	4.02	3.05	2.71	2.54	2.41	2.27	2.21	2.12	2.08	2.04	2.02
	0.01	113.4	7.82	5.15	4.11	3.71	3.44	3.23	3.05	2.86	2.81	2.71	2.66
1.8	0.10	11.13	3.00	2.37	2.14	2.03	1.95	1.87	1.82	1.75	1.73	1.69	1.68
	0.05	22.60	4.05	3.05	2.70	2.53	2.40	2.27	2.22	2.12	2.08	2.04	2.02
	0.01	117.2	7.94	5.10	4.10	3.67	3.42	3.24	3.05	2.86	2.81	2.71	2.67
1.9	0.10	11.10	3.00	2.37	2.14	2.03	1.95	1.87	1.82	1.75	1.73	1.69	1.68
	0.05	22.71	4.04	3.05	2.71	2.53	2.40	2.27	2.22	2.12	2.08	2.04	2.02
	0.01	121.4	7.92	5.01	4.10	3.68	3.43	3.23	3.04	2.86	2.81	2.71	2.67
2	0.10	11.05	3.01	2.37	2.14	2.03	1.95	1.87	1.82	1.75	1.73	1.69	1.67
	0.05	22.56	4.06	3.06	2.71	2.53	2.41	2.27	2.22	2.12	2.07	2.04	2.02
	0.01	115.7	7.89	4.94	4.09	3.65	3.42	3.21	3.04	2.86	2.81	2.71	2.67

TABLE 2. 5% EMPIRICAL SIZES OF $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1,2})$

			m											
			1		2		3		5		10		\sqrt{n}	
γ	δ	n	F	N	F	N	F	N	F	N	F	N	F	N
0.2	0.4	64	.047	.460	.061	.252	.072	.212	.122	.227	.279	.355	.215	.296
		128	.049	.462	.053	.235	.057	.180	.079	.161	.189	.256	.219	.283
		256	.050	.460	.050	.224	.049	.164	.059	.127	.112	.157	.218	.263
		512	.048	.456	.052	.234	.052	.165	.054	.115	.064	.101	.204	.220
0.0	0.4	64	.049	.459	.060	.253	.068	.202	.105	.200	.221	.286	.170	.244
		128	.048	.460	.053	.232	.056	.174	.071	.147	.148	.201	.164	.217
		256	.046	.461	.052	.222	.049	.163	.059	.122	.097	.135	.166	.197
		512	.052	.456	.053	.233	.052	.163	.054	.114	.062	.095	.154	.165
0.4	1.2	64	.050	.461	.055	.238	.057	.179	.073	.154	.123	.173	.102	.159
		128	.052	.459	.053	.225	.051	.166	.060	.126	.084	.122	.091	.127
		256	.048	.462	.053	.224	.048	.160	.052	.115	.064	.098	.086	.110
		512	.047	.457	.054	.230	.050	.164	.050	.112	.055	.087	.080	.088
0.0	1.0	64	.051	.463	.054	.235	.052	.171	.063	.137	.095	.136	.083	.128
		128	.050	.455	.052	.225	.051	.160	.055	.119	.066	.098	.068	.097
		256	.050	.463	.050	.221	.048	.161	.051	.113	.056	.086	.064	.085
		512	.045	.458	.054	.228	.051	.163	.049	.109	.053	.080	.064	.072
0.4	2.0	64	.050	.460	.052	.222	.046	.155	.046	.107	.056	.087	.052	.089
		128	.050	.458	.049	.220	.044	.152	.045	.103	.044	.070	.046	.068
		256	.049	.457	.050	.219	.049	.151	.045	.104	.044	.068	.042	.056
		512	.050	.454	.052	.226	.047	.158	.047	.107	.047	.076	.050	.056

TABLE 3. 5% EMPIRICAL SIZES OF $\hat{t}_m(\hat{\gamma}, \hat{\delta}, \hat{\omega}_{1,2})$

			m											
			1		2		3		5		10		\sqrt{n}	
γ	δ	n	F	N	F	N	F	N	F	N	F	N	F	N
0.2	0.4	64	.072	.570	.160	.421	.211	.388	.274	.380	.340	.393	.319	.384
		128	.069	.562	.162	.424	.223	.403	.299	.410	.367	.419	.374	.416
		256	.068	.555	.142	.411	.207	.401	.306	.429	.399	.457	.426	.460
		512	.064	.529	.124	.374	.178	.360	.274	.394	.402	.460	.483	.499
0.0	0.4	64	.067	.554	.138	.392	.176	.348	.226	.331	.281	.331	.266	.327
		128	.064	.549	.136	.394	.187	.363	.247	.360	.308	.362	.312	.357
		256	.062	.540	.126	.381	.177	.364	.252	.377	.334	.393	.358	.396
		512	.063	.520	.112	.356	.151	.331	.223	.343	.333	.393	.406	.421
0.4	1.2	64	.059	.510	.091	.310	.109	.253	.132	.216	.160	.203	.150	.203
		128	.057	.502	.087	.309	.112	.260	.134	.225	.159	.205	.162	.204
		256	.053	.501	.083	.303	.105	.257	.133	.229	.167	.216	.174	.204
		512	.053	.490	.081	.288	.095	.240	.119	.209	.153	.201	.182	.194
0.0	1.0	64	.052	.487	.074	.272	.083	.210	.098	.172	.121	.156	.115	.158
		128	.051	.481	.070	.273	.083	.213	.095	.171	.113	.150	.117	.148
		256	.056	.487	.073	.274	.084	.216	.098	.179	.115	.155	.119	.145
		512	.051	.479	.069	.268	.078	.212	.091	.171	.111	.149	.124	.135
0.4	2.0	64	.043	.432	.054	.210	.062	.155	.071	.124	.089	.116	.084	.117
		128	.045	.434	.047	.211	.053	.151	.060	.117	.071	.097	.075	.098
		256	.047	.445	.049	.215	.053	.151	.056	.109	.063	.089	.070	.084
		512	.047	.450	.056	.228	.052	.158	.051	.113	.059	.087	.074	.081

TABLE 4. 10% EMPIRICAL SIZES OF $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$

			m											
			1		2		3		5		10		\sqrt{n}	
γ	δ	n	F	N	F	N	F	N	F	N	F	N	F	N
0.2	0.4	64	.096	.517	.116	.323	.144	.283	.215	.311	.400	.461	.324	.400
		128	.098	.518	.104	.301	.115	.246	.149	.232	.297	.355	.332	.382
		256	.099	.517	.097	.290	.104	.227	.118	.195	.189	.235	.328	.366
		512	.097	.512	.103	.298	.106	.231	.106	.178	.130	.170	.311	.325
0.0	0.4	64	.097	.515	.114	.316	.135	.272	.187	.281	.328	.385	.269	.339
		128	.097	.515	.105	.297	.114	.241	.137	.217	.239	.293	.266	.313
		256	.096	.517	.101	.289	.103	.225	.113	.187	.163	.206	.258	.291
		512	.099	.515	.103	.299	.106	.229	.105	.176	.119	.161	.244	.254
0.4	1.2	64	.100	.517	.107	.304	.116	.248	.141	.227	.209	.256	.179	.237
		128	.103	.514	.100	.293	.109	.230	.117	.190	.154	.194	.159	.200
		256	.097	.518	.099	.288	.102	.221	.106	.178	.124	.160	.151	.172
		512	.096	.512	.104	.293	.107	.227	.102	.175	.111	.145	.146	.153
0.0	1.0	64	.098	.516	.104	.302	.113	.238	.126	.205	.167	.213	.147	.205
		128	.098	.515	.099	.291	.107	.226	.109	.176	.123	.161	.127	.161
		256	.099	.515	.098	.291	.102	.219	.103	.174	.110	.145	.120	.140
		512	.095	.512	.104	.293	.104	.227	.099	.173	.103	.139	.121	.127
0.4	2.0	64	.099	.516	.099	.286	.097	.217	.098	.169	.109	.144	.101	.146
		128	.099	.512	.097	.284	.097	.220	.094	.164	.091	.123	.091	.120
		256	.101	.516	.095	.289	.097	.216	.095	.163	.087	.121	.088	.106
		512	.101	.512	.099	.292	.099	.222	.099	.174	.096	.130	.104	.109

TABLE 5. 10% EMPIRICAL SIZES OF $\hat{t}_m(\hat{\gamma}, \hat{\delta}, \hat{\omega}_{1.2})$

			m											
			1		2		3		5		10		\sqrt{n}	
γ	δ	n	F	N	F	N	F	N	F	N	F	N	F	N
0.2	0.4	64	.139	.622	.247	.487	.309	.458	.367	.457	.430	.470	.405	.458
		128	.135	.616	.253	.495	.323	.481	.397	.486	.452	.493	.455	.491
		256	.132	.609	.234	.489	.317	.480	.416	.508	.493	.537	.508	.536
		512	.126	.585	.205	.448	.277	.440	.380	.475	.498	.542	.560	.568
0.0	0.4	64	.131	.609	.221	.460	.273	.424	.318	.408	.365	.410	.347	.406
		128	.125	.602	.224	.465	.286	.442	.345	.442	.395	.439	.399	.437
		256	.128	.599	.209	.455	.282	.443	.361	.462	.435	.481	.448	.480
		512	.123	.577	.188	.424	.248	.407	.329	.428	.433	.479	.494	.504
0.4	1.2	64	.119	.563	.161	.380	.188	.328	.203	.292	.232	.272	.221	.273
		128	.111	.559	.158	.381	.190	.335	.212	.304	.235	.277	.237	.276
		256	.113	.556	.151	.375	.186	.334	.218	.309	.246	.293	.254	.280
		512	.107	.545	.143	.358	.168	.314	.196	.287	.239	.286	.267	.275
0.0	1.0	64	.108	.539	.134	.340	.151	.279	.163	.236	.181	.221	.173	.220
		128	.102	.534	.133	.338	.152	.285	.163	.242	.176	.214	.177	.211
		256	.112	.544	.132	.342	.150	.290	.165	.251	.182	.226	.189	.208
		512	.101	.531	.130	.335	.146	.279	.161	.241	.175	.219	.196	.203
0.4	2.0	64	.092	.487	.102	.269	.109	.207	.115	.177	.133	.162	.129	.161
		128	.091	.488	.094	.268	.104	.206	.108	.167	.117	.145	.121	.144
		256	.095	.499	.098	.277	.102	.209	.101	.164	.107	.140	.118	.136
		512	.096	.504	.104	.288	.101	.221	.104	.172	.107	.141	.125	.132

TABLE 6. 5% EMPIRICAL SIZES OF $\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$ AND $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$

n	CN ($m = \lfloor n^{0.4} \rfloor$)			F ($m = 3$)			F ($m = 10$)		
	I	II	III	I	II	III	I	II	III
64	.172	.230	.527	.046	.045	.068	.046	.060	.221
128	.165	.187	.604	.046	.049	.056	.049	.050	.148
256	.151	.161	.761	.047	.048	.049	.049	.053	.097
512	.139	.143	.867	.050	.051	.052	.049	.049	.060

TABLE 7. 5% EMPIRICAL POWER OF $\tilde{\tau}_n, \tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$ AND $\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$

c	0	20	50	100	150	200	300	500
$\tilde{\tau}_n$	0.05	0.22	0.81	1.00	1.00	1.00	1.00	1.00
$\tilde{t}_m(\gamma, \delta, \hat{\omega}_{11}, \hat{\omega}_{22})$ ($m=10$)	0.05	0.09	0.34	0.91	1.00	1.00	1.00	1.00
$\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$ ($m=10$)	0.05	0.07	0.20	0.57	0.85	0.96	1.00	1.00
$\hat{t}_m(\gamma, \delta, \hat{\omega}_{1.2})$ ($m=3$)	0.05	0.06	0.10	0.28	0.51	0.69	0.90	0.99

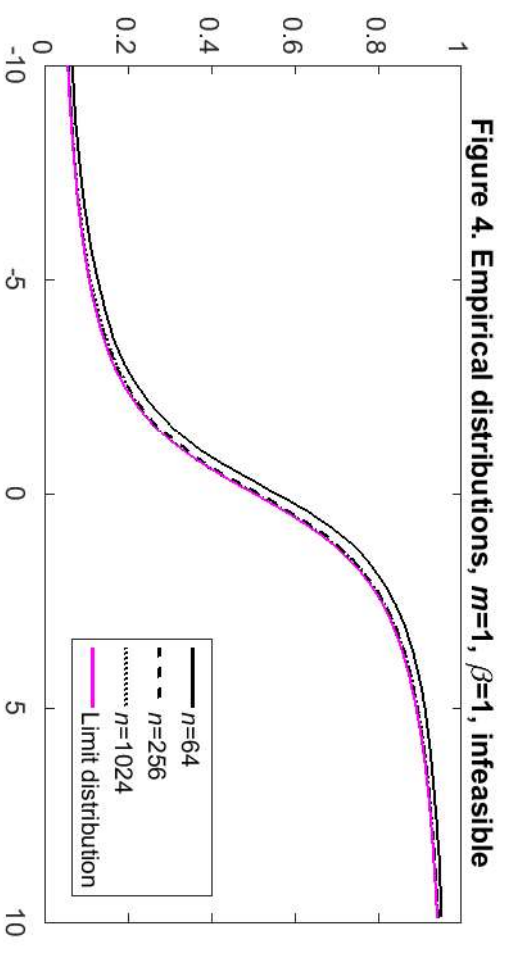
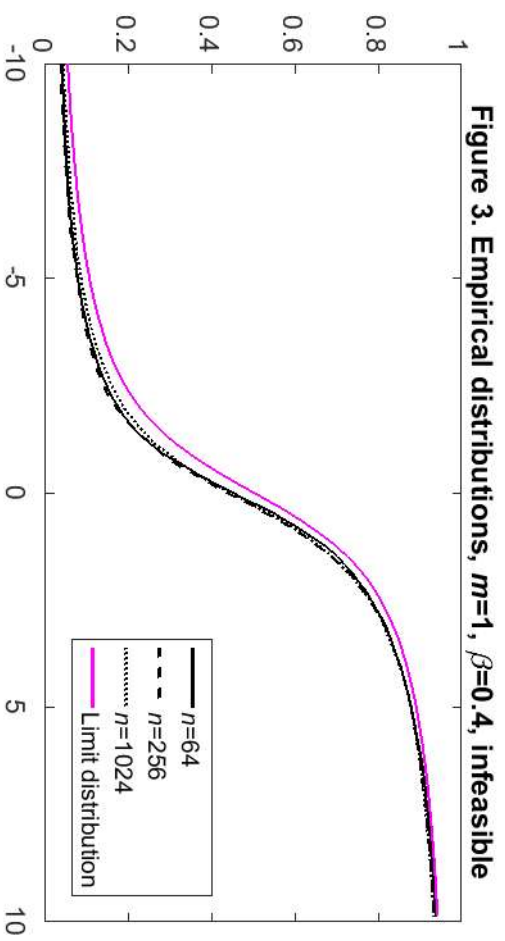
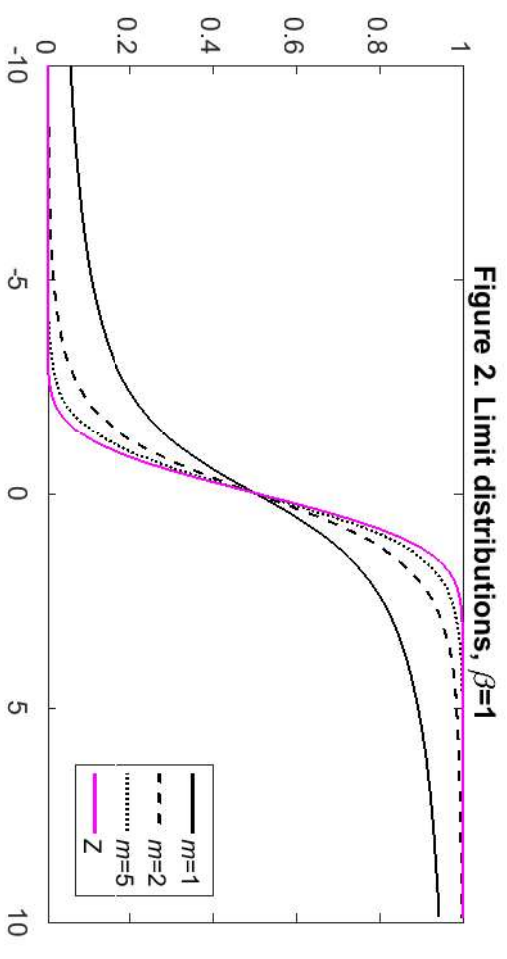
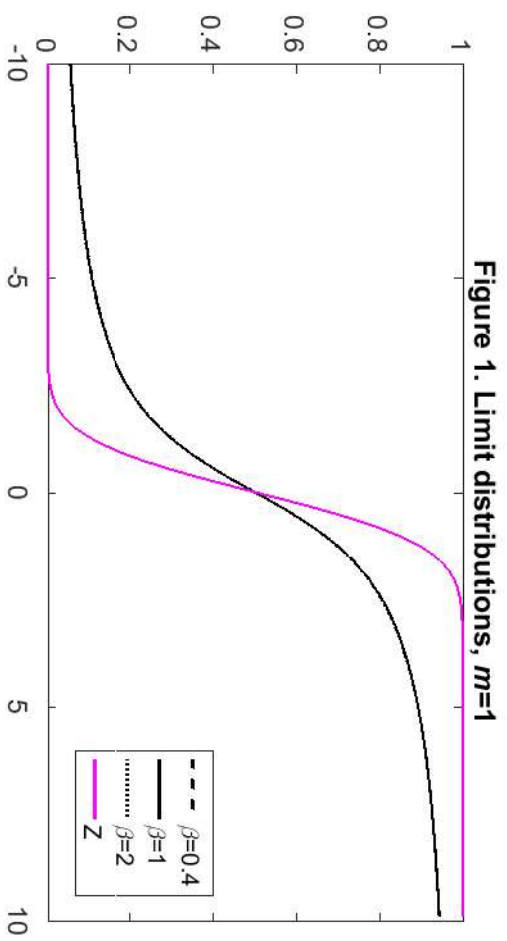


Figure 5. Empirical distributions, $m=2$, $\beta=1$, infeasible

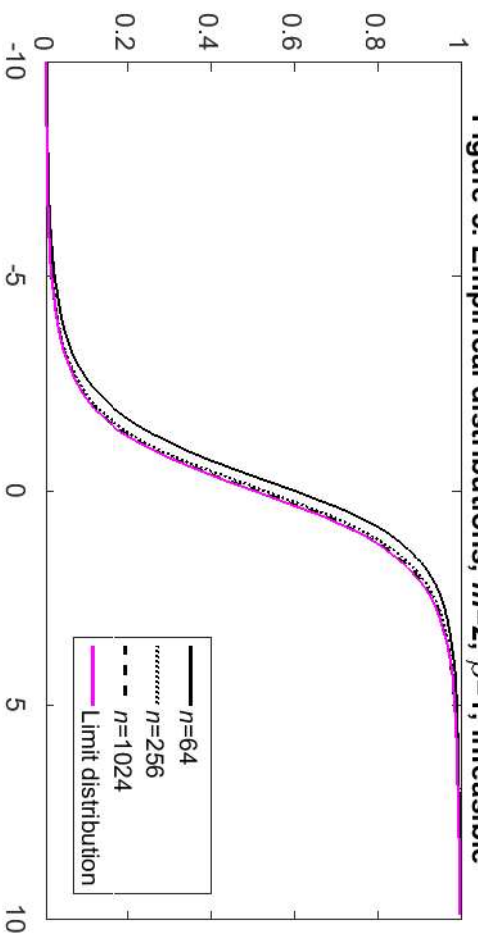


Figure 6. Empirical distributions, $m=3$, $\beta=1$, infeasible

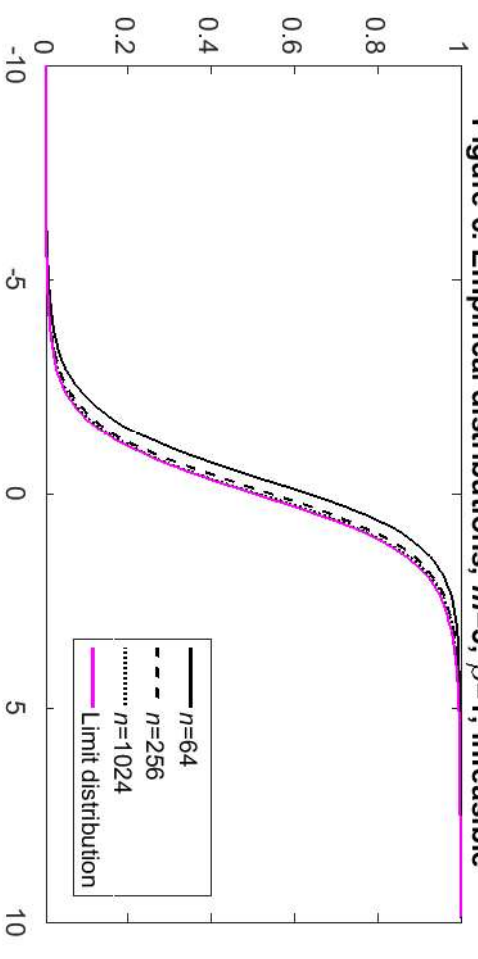


Figure 7. Empirical distributions, $m=1$, $\beta=0.4$, feasible

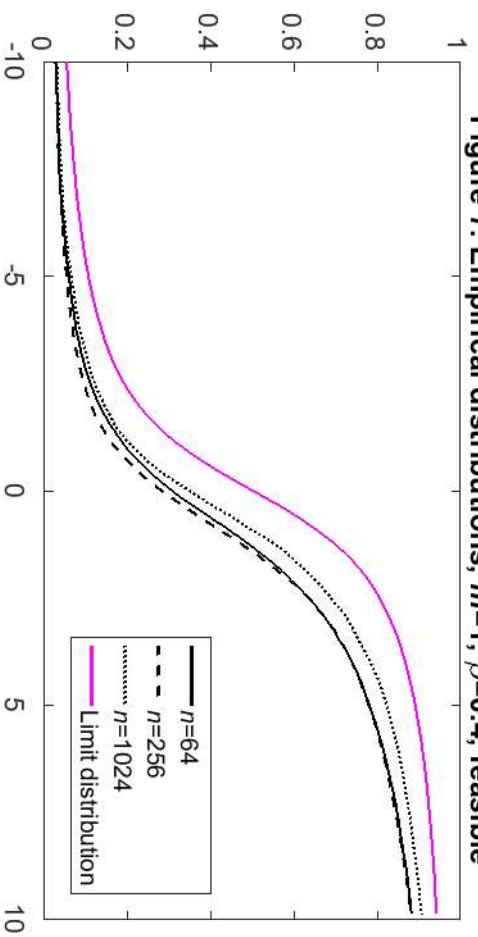


Figure 8. Empirical distributions, $m=1$, $\beta=1$, feasible

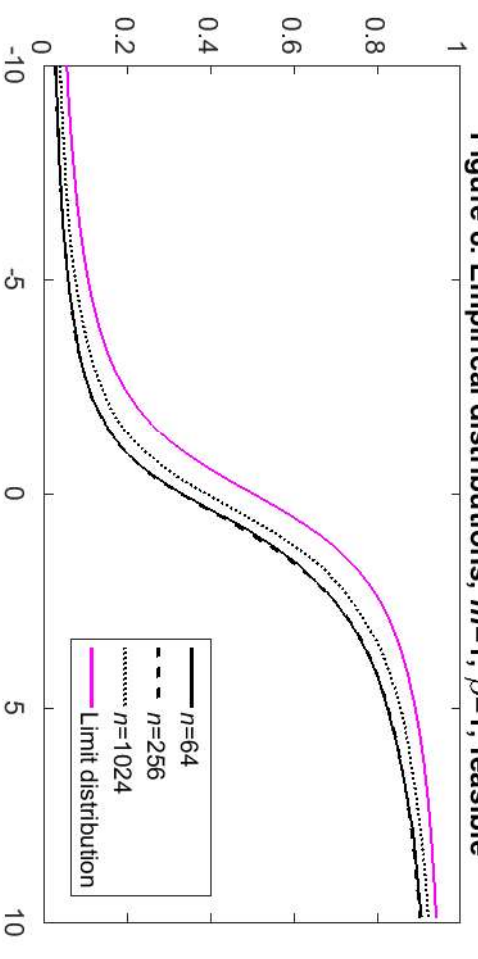


Figure 9. Empirical distributions, $m=2$, $\beta=1$, feasible

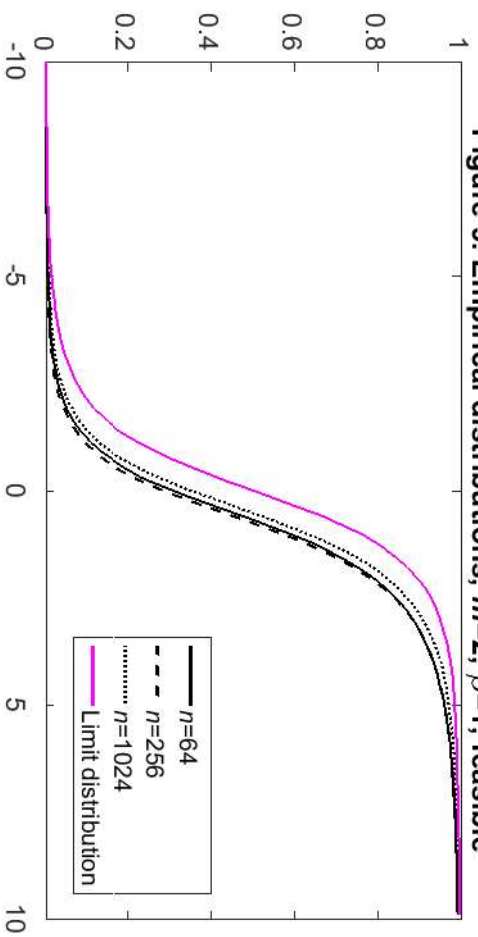


Figure 10. Empirical distributions, $m=3$, $\beta=1$, feasible

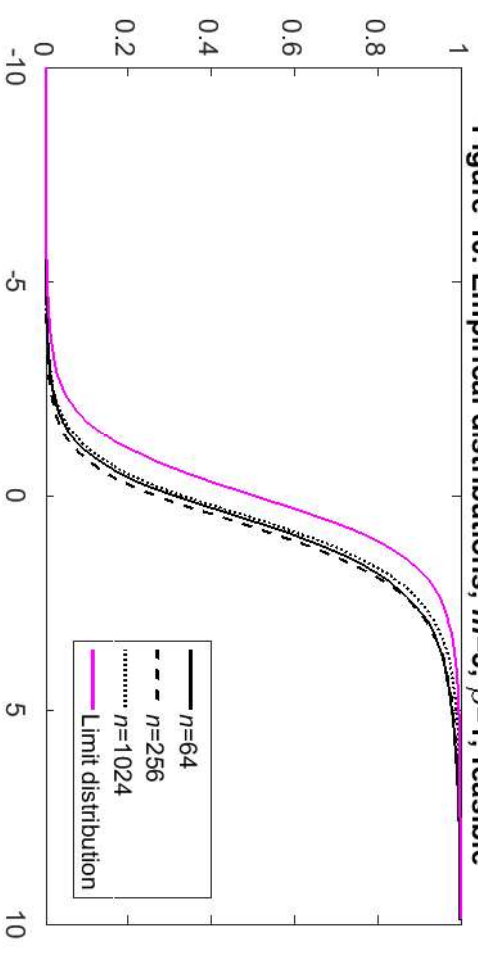


Figure 11. Empirical distributions, $m=10$, $\beta=1$, feasible

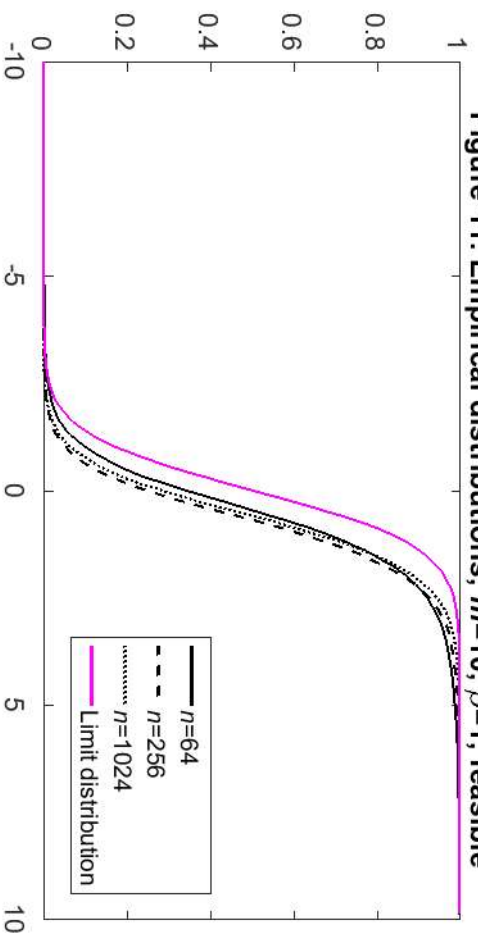


Figure 12. Local Power at 0.05 level, $n=512$, $\beta=1$

